Section 45: EX-32.A (1350 CERTIFICATION OF CEO - SWEPCO)

Exhibit 32(a)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63
of Title 18 of the United States Code

In connection with the Annual Report of Southwestern Electric Power Company (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Nicholas K. Akins, the chief executive officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Nicholas K. Akins
Nicholas K. Akins
Chief Executive Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to Southwestern Electric Power Company and will be retained by Southwestern Electric Power Company and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 46: EX-32.B (1350 CERTIFICATION OF CFO - AEP)

Exhibit 32(b)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63 of Title 18 of the United States Code

In connection with the Annual Report of American Electric Power Company, Inc. (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Brian X. Tierney, the chief financial officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Brian X. Tierney
Brian X. Tierney
Chief Financial Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to American Electric Power Company, Inc. and will be retained by American Electric Power Company, Inc. and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 47: EX-32.B (1350 CERTIFICATION OF CFO - AEPTCO)

Exhibit 32(b)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63
of Title 18 of the United States Code

In connection with the Annual Report of AEP Transmission Company, LLC (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Brian X. Tierney, the chief financial officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Brian X. Tierney
Brian X. Tierney
Chief Financial Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to AEP Transmission Company, LLC and will be retained by AEP Transmission Company, LLC and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 48: EX-32.B (1350 CERTIFICATION OF CFO - AEP TEXAS)

Exhibit 32(b)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63
of Title 18 of the United States Code

In connection with the Annual Report of AEP Texas Inc. (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Brian X. Tierney, the chief financial officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Brian X. Tierney
Brian X. Tierney
Chief Financial Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to AEP Texas Inc. and will be retained by AEP Texas Inc. and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 49: EX-32.B (1350 CERTIFICATION OF CFO - APCO)

Exhibit 32(b)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63
of Title 18 of the United States Code

In connection with the Annual Report of Appalachian Power Company (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Brian X. Tierney, the chief financial officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Brian X. Tierney
Brian X. Tierney
Chief Financial Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to Appalachian Power Company and will be retained by Appalachian Power Company and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 50: EX-32.B (1350 CERTIFICATION OF CFO - I&M)

Exhibit 32(b)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63
of Title 18 of the United States Code

In connection with the Annual Report of Indiana Michigan Power Company (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Brian X. Tierney, the chief financial officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Brian X. Tierney
Brian X. Tierney
Chief Financial Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to Indiana Michigan Power Company and will be retained by Indiana Michigan Power Company and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 51: EX-32.B (1350 CERTIFICATION OF CFO - OPCO)

Exhibit 32(b)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63
of Title 18 of the United States Code

In connection with the Annual Report of Ohio Power Company (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Brian X. Tierney, the chief financial officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Brian X. Tierney

Brian X. Tierney
Chief Financial Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to Ohio Power Company and will be retained by Ohio Power Company and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 52: EX-32.B (1350 CERTIFICATION OF CFO - PSO)

Exhibit 32(b)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63
of Title 18 of the United States Code

In connection with the Annual Report of Public Service Company of Oklahoma (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Brian X. Tierney, the chief financial officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Brian X. Tierney
Brian X. Tierney
Chief Financial Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to Public Service Company of Oklahoma and will be retained by Public Service Company of Oklahoma and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 53: EX-32.B (1350 CERTIFICATION OF CFO - SWEPCO)

Exhibit 32(b)

This Certification is being furnished and shall not be deemed "filed" for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liability of that section. This Certification shall not be incorporated by reference into any registration statement or other document pursuant to the Securities Act of 1933, except as otherwise stated in such filing.

Certification Pursuant to Section 1350 of Chapter 63
of Title 18 of the United States Code

In connection with the Annual Report of Southwestern Electric Power Company (the "Company") on Form 10-K (the "Report") for the year ended December 31, 2019 as filed with the Securities and Exchange Commission on the date hereof, I, Brian X. Tierney, the chief financial officer of the Company certify pursuant to 18 U.S.C. Section 1350, as adopted pursuant to Section 906 of the Sarbanes-Oxley Act of 2002 that, based on my knowledge (i) the Report fully complies with the requirements of Section 13(a) or 15(d) of the Securities Exchange Act of 1934 and (ii) the information contained in the Report fairly presents, in all material respects, the financial condition and results of operations of the Company.

/s/ Brian X. Tierney
Brian X. Tierney
Chief Financial Officer

February 20, 2020

A signed original of this written statement required by Section 906 has been provided to Southwestern Electric Power Company and will be retained by Southwestern Electric Power Company and furnished to the Securities and Exchange Commission or its staff upon request.

(Back To Top)

Section 54: EX-32.B (MINE SAFETY DISCLOSURE)

Exhibit 95

MINE SAFETY INFORMATION

The Federal Mine Safety and Health Act of 1977 (Mine Act) imposes stringent health and safety standards on various mining operations. The Mine Act and its related regulations affect numerous aspects of mining operations, including training of mine personnel, mining procedures, equipment used in mine emergency procedures, mine plans and other matters. SWEPCo, through its ownership of Dolet Hills Lignite Company (DHLC), a wholly-owned lignite mining subsidiary of SWEPCo, is subject to the provisions of the Mine Act.

The Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act) requires companies that operate mines to include in their periodic reports filed with the SEC, certain mine safety information covered by the Mine Act. DHLC received the following notices of violation and proposed assessments under the Mine Act for the quarter-ended December 31, 2019:

Number of Citations for S&S Violations of Mandatory Health or Safety Standards under 104 *		0
Number of Orders Issued under 104(b) *		0
Number of Citations and Orders for Unwarrantable Failure to Comply with Mandatory Health or Safety Standards under 104(d) *		0
Number of Flagrant Violations under 110(b)(2) *		0
Number of Imminent Danger Orders Issued under 107(a)		0
Total Dollar Value of Proposed Assessments **	$	—
Number of Mining-related Fatalities		0

* References to sections under the Mine Act.
** DHLC received two non-S&S citations during the fourth quarter of 2019. Proposed assessments for those citations were not received in 2019.

There are currently no legal actions pending before the Federal Mine Safety and Health Review Commission.

(Back To Top)

FUNDAMENTALS OF FINANCIAL MANAGEMENT

CONCISE FOURTH EDITION

EUGENE F. BRIGHAM
University of Florida

JOEL F. HOUSTON
University of Florida

THOMSON
SOUTH-WESTERN

Maturity Matching or "Self-Liquidating," Approach

Maturity Matching, or "Self-Liquidating," Approach
A financing policy that matches asset and liability maturities. This is a moderate policy.

The **maturity matching,** or **"self-liquidating," approach** calls for matching asset and liability maturities as shown in Panel a of Figure 14-3. This strategy minimizes the risk that the firm will be unable to pay off its maturing obligations. To illustrate, suppose a company borrows on a one-year basis and uses the funds obtained to build and equip a plant. Cash flows from the plant (profits plus depreciation) would not be sufficient to pay off the loan at the end of only one year, so the loan would have to be renewed. If for some reason the lender refused to renew the loan, then the company would have problems. Had the plant been financed with long-term debt, however, the required loan payments would have been better matched with cash flows from profits and depreciation, and the problem of renewal would not have arisen.

At the limit, a firm could attempt to match exactly the maturity structure of its assets and liabilities. Inventory expected to be sold in 30 days could be financed with a 30-day bank loan; a machine expected to last for 5 years could be financed with a 5-year loan; a 20-year building could be financed with a 20-year mortgage bond; and so forth. Actually, of course, two factors prevent this exact maturity matching: (1) there is uncertainty about the lives of assets, and (2) some common equity must be used, and common equity has no maturity. To illustrate the uncertainty factor, a firm might finance inventories with a 30-day loan, expecting to sell the inventories and then use the cash to retire the loan. But if sales were slow, the cash would not be forthcoming, and the use of short-term credit could end up causing a problem. Still, if a firm makes an attempt to match asset and liability maturities, we would define this as a moderate current asset financing policy.

In practice, firms don't finance each specific asset with a type of capital that has a maturity equal to the asset's life. However, academic studies do show that most firms tend to finance short-term assets from short-term sources and long-term assets from long-term sources.[16]

Aggressive Approach

WWW Students can access various types of historical interest rates, including fixed and variable rates, at the St. Louis Federal Reserve's FRED site. The address is **http://research.stlouisfed.org/fred/**.

Panel b of Figure 14-3 illustrates the situation for a relatively aggressive firm that finances all of its fixed assets with long-term capital and part of its permanent current assets with short-term, nonspontaneous credit. Note that we used the term "relatively" in the title for Panel b because there can be different *degrees* of aggressiveness. For example, the dashed line in Panel b could have been drawn *below* the line designating fixed assets, indicating that all of the permanent current assets and part of the fixed assets were financed with short-term credit; this would be a highly aggressive, extremely nonconservative position, and the firm would be very much subject to dangers from rising interest rates as well as to loan renewal problems. However, short-term debt is often cheaper than long-term debt, and some firms are willing to sacrifice safety for the chance of higher profits.

Conservative Approach

Panel c of Figure 14-3 has the dashed line *above* the line designating permanent current assets, indicating that long-term capital is being used to finance all permanent

[16] For example, see William Beranek, Christopher Cornwell, and Sunho Choi, "External Financing, Liquidity, and Capital Expenditures," *Journal of Financial Research*, Vol. 18, No. 2, 207–222.

J Regul Econ (2011) 40 261–278
DOI 10.1007/s11149-011-9160-5

ORIGINAL ARTICLE

New approach to estimating the cost of common equity capital for public utilities

Pauline M. Ahern · Frank J. Hanley · Richard A. Michelfelder

Published online 26 August 2011

Abstract The regulatory process for setting public utilities' allowed rate of return on common equity has generally used the Gordon DCF, CAPM and Risk Premium specifications to estimate the cost of common equity. Despite the widely known problems with these models, there has been little movement to adopt more recently developed asset pricing models to provide additional evidence for estimating the cost of capital. This paper presents, validates empirically and applies a general yet simple consumption-based asset pricing specification to model the risk-return relationship for stocks and estimate the cost of common equity for public utilities. The model is not necessarily superior to other models in its practical results, yet these results do indicate that it should be used to provide additional estimates of the cost of common equity. Additionally, the model raises doubts as to whether assets such as utility stocks are a consumption (business cycle) hedge.

Keywords Public utilities · Cost of capital · GARCH Consumption asset pricing model

JEL Classification G12 · L94 · L95

P. M. Ahern · F. J. Hanley
AUS Consultants, Mt. Laurel, NJ 08054, USA
e-mail: pahern@ausinc.com

F. J. Hanley
e-mail: Fhanley@ausinc.com

R. A. Michelfelder (✉)
School of Business, Rutgers University,
Camden, NJ 08102-1656, USA
e-mail: richmich@rutgers.edu

1 Introduction

Following electricity deregulation with the National Energy Policy Act of 1992, the estimation of the cost of common equity capital remains a critical component of the utility rate-of-return regulatory process. Since the cost of common equity is not observable in capital markets, it must be inferred from asset pricing models. The models that are commonly applied in regulatory proceedings are the Gordon (1974) Discounted Cash Flow (DCF), the Capital Asset Pricing (CAPM) and Risk Premium Models. There are other tools used to estimate the cost of common equity such as comparable earnings or earnings-to-price ratios, but they are not asset pricing models. The empirical literature on the CAPM is vast {Fama and French (2004)} and the CAPM is used by a number of US regulatory jurisdictions. The DCF model has not been empirically tested to the same extent as the CAPM, yet it is considered by many US regulatory jurisdictions

The purpose of this paper is to present, test empirically and apply a recently developed general consumption-based asset pricing model that estimates the risk-return relationship directly from asset pricing data and, when estimated with recently developed time series methods, produces a prediction of the equity risk premium that is driven by its predicted volatility. The predicted risk premium is then added to a risk-free rate of return to provide an estimate of the cost of common equity. We predict two forms of the equity risk premium with the model, the risk premium net of the risk-free rate and the equity-to-debt risk premium (equity risk premium net of the relevant bond yield for the company's stock). Either can be applied to predict the common equity cost of capital for a public utility. Although the model is tested and applied to public utilities for rate of return regulation, it can be used to estimate the cost of capital for any stock. Section 2 reviews the asset pricing models typically used in public utility rate cases and the generalized consumption asset pricing model we propose to estimate the cost of common equity. Section 3 discusses the data and the empirical testing of the consumption asset pricing model Section 4 reviews the application of the model and compares it with the DCF and CAPM results. Section 5 is the conclusion.

2 DCF, CAPM and consumption asset pricing model

2.1 DCF and CAPM approaches

The standard DCF model frequently used in estimative the cost rate of common equity in regulatory proceedings is defined by the following equation:

$$k = D_0(1+g)/P_0 + g,$$

where k is the expected return on common equity, D_0 is the current dividend per share, g is the expected dividend per share growth rate; and P_0 is the current market price.

The DCF was developed by Gordon (1974) specifically for regulatory purposes. Underlying the DCF model is the theory that the present value of an expected future stream of net cash flows during the investment holding period can be determined by discounting those cash flows at the cost of capital, or the investors' capitalization rate. DCF theory indicates that an investor buys a stock for an expected total return rate which is derived from cash flows received in the form of dividends plus appreciation in market price (the expected growth rate) over the investment holding period Mathematically, the expected dividend yield $(D_0(1+g)/P_0)$ on market price plus an expected growth rate equals the capitalization rate, i.e., the expected return on common equity.

The standard DCF contains several restrictive assumptions, the most contentious of which during utility cost of capital proceedings is typically that dividends per share (DPS), book value per share (BVPS), earnings per share (EPS) as well as market price grow at the same rate in perpetuity. There is also considerable contention over the proper proxy for g, prospective or historical growth in DPS, BVPS, EPS and market price and over what time period. In addition, although the standard DCF described above is a single stage annual growth model, there is considerable discussion over the use of multiple stage growth models during regulatory proceedings. Some analysts use the discrete version and others use the continuous version of the DCF model Solving these models for k, the cost of common equity, results in differing equations to solve for k. The equation above is from the discrete version. The continuous version uses the current dividend yield and is not adjusted by g, which results in a lower estimate for k Because of these and other restrictive assumptions that require numerous subjective judgments in application, it is often difficult for regulatory commissions to reconcile the frequently large disparities in rates of return on common equity recommended by various parties in a public utility rate case.

The CAPM model is defined by the following equation:

$$k = R_f + \beta\left(R_m - R_f\right),$$

where k is the expected return on common equity; R_f is the expected risk-free rate of return; β is the expected beta; and R_m is the expected market return.

CAPM theory defines risk as the co-variability of a security's returns with the market's returns or β, also known as systematic or market risk, with the market beta being defined as 1.0. Because CAPM theory assumes that all investors hold perfectly diversified portfolios, they are presumed to be exposed only to systematic risk and the market (according to the model) will not reward them a risk premium for unsystematic or non-market risk. In other words, the CAPM presumes that investors require compensation only for systematic or market risks which are due to macroeconomic and other events that affect the returns on all assets. Mathematically, the CAPM is applied by adding a forward-looking risk-free rate of return to an expected market equity risk premium adjusted proportionately by the expected beta to reflect the systematic risk.

As with the DCF, there is considerable contention during regulatory cost of capital proceedings as to the proper proxies for all components of the CAPM: the R_f, the R_m, as well as β In addition, the CAPM assumption that the market will only reward investors for systematic or market risk is extremely restrictive when estimating the expected return on common equity for a single asset such as a single jurisdictional regulated operating utility. Additionally, this assumption requires that the investor have a perfectly diversified portfolio, that is, one with no unsystematic risk. Since this assumption is not applicable, estimating the cost of common equity capital for a single utility's common equity undoubtedly will not reflect the risk actually faced by the imperfectly diversified investor.

As will be discussed in the next section, our application of the risk premium approach, the consumption asset pricing model and GARCH[1] rest on minimal assumptions and restrictions and therefore requires considerably less judgment in its application.

2.2 Risk premium approach, consumption asset pricing models, and GARCH

A widely used model to estimate the cost of common equity capital for public utilities is the risk premium approach. This approach often estimates the expected rate of return as the long-term historic mean of the realized risk premium above an historic yield plus the current yield of the relevant bond applicable to a specific utility or peer group of utilities. Litigants in public utility rate proceedings debate the choice of inputs to estimate the risk premium as well as how far back to reach into history to collect data for calculating an average that is representative of a forward-looking premium.

It is surprising that, as popular as the risk premium method is in public utility rate cases, the intuitively appealing general consumption-based asset pricing model, with its minimal assumptions and strong theoretical foundation, has not been applied to estimate the cost of common equity capital for public utilities. The model provides projections of the conditional expected risk premium on an asset based on its relation to its predicted conditional volatility. This model generalizes the well known special case asset pricing models such as the Merton (1973) intertemporal capital asset pricing model, Campbell (1993) intertemporal asset pricing model, and the habit-persistence model of Campbell and Cochrane (1999), which are special cases of the general model. The relation of the model to their specialized cases can be found in Cochrane (2006) and Cochrane (2007). The approach of consumption asset pricing models is to make investment decisions that maximize investors' utility from the consumption that they ultimately desire, not returns.

Even if the model is not used to project directly the expected risk premium, it can, at a minimum, be used to verify that the risk premia data chosen for estimating the cost of capital is empirically validated by fitting the model well. The model can be used to predict the equity risk premia net of the risk-free rate (equity risk premium) or to predict the equity-to-debt risk premium for a firm. We perform both of these empirical tests in this paper. The general consumption-based asset pricing model developed in Michelfelder and Pilotte (2011) and based on Cochrane (2004) provides the relationship of the ex ante risk premium to an asset's own volatility in return:

$$E_t[R_{i,t+1}] - R_{f,t} = -\frac{vol_t[M_{t+1}]}{E_t\left[M_{t+1}\right]} vol_t[R_{i,t+1}] corr_t[M_{t+1}, R_{i,t+1}]. \tag{1}$$

[1] GARCH refers to the generalized autoregressive conditional heteroskedasticity regression model which is discussed below

where vol_t is the conditional volatility, $corr_t$ is the conditional correlation, and M_{t+1} is the stochastic discount factor (SDF).

The SDF is the intertemporal marginal rate of substitution in consumption, or, $M_{t+1} = \beta \frac{U_{c,t+1}}{U_{c,t}}$, where the U_c's are the marginal utilities of consumption in the next period, $t+1$, and the current period, t, and β is the discount factor for period t to $t+1$. Equation 1 shows that the algebraic sign of the relation between the expected risk premium and the conditional volatility of an asset's risk premium is determined by the correlation between the asset's return and the SDF That is, the direction of the relation between the asset return and the ratio of intertemporal marginal utilities in consumption inversely determines the relation between the expected risk premium and conditional volatility. When the correlation is equal to negative one, the asset's conditional expected risk premium is perfectly positively correlated with its conditional volatility. A positive relation between the conditionally expected risk premium and volatility obtains when $-1 < corr_t < 0$. A negative relation obtains when $0 < corr_t < 1$. For an asset that represents a perfect hedge against shocks to the marginal utility of consumption, with $corr_t = 1$, there will be a perfect negative correlation between the conditionally expected risk premium and its volatility.[2] Therefore, estimates of the relation between the first two conditional moments of a public utility stock's returns provide a direct test of the effectiveness of a public utility stock, or any asset, as a consumption hedging asset. In Eq. 1, $vol_t[M_{t+1}]/E_t[M_{t+1}]$ is the slope of the mean-variance frontier. If this slope changes over time, the estimated relation between the stock's risk and return will vary over time. This model can also be viewed simplistically as the projected expected risk premium as a function of its own projected risk, given information available at time t.

Note that the model allows for the expected risk premium to be negative if the asset hedges shocks to the marginal utility of consumption. Investors are willing to accept an expected rate of return lower than the risk-free rate of return if the pattern of volatility is such that returns are expected to rise with expected reductions in consumption Simply, investors are willing to *pay* a premium for a higher level of returns volatility that has the desired pattern of returns. These desired returns patterns have a tendency to offset drops in consumption Therefore, this model shows that investors may not be averse to volatility, but rather to the timing of expected changes in returns.

Summarizing, several conclusions can be drawn from the general model of asset pricing. First, the sign of the relation between a stock's risk premium and conditional volatility depends on the extent to which the stock serves as an intertemporal hedge against shocks to the marginal utility of consumption. Second, the relation between stock risk and return may be time-varying depending on changes in the slope of the mean-variance frontier Third, hedging assets have desired patterns of volatility that result in expected rates of return that are less than the risk-free rate. We do not expect

[2] A hedging asset is one that has a positive increase in returns that is coincident with a positive shock in the ratio of intertemporal marginal utilities of consumption Note that if we assume a concave utility function in consumption, as consumption declines, the marginal utility of consumption rises relative to last period marginal utility If we think of a decline in consumption as a contraction in the business cycle, the hedging asset delivers positive changes in returns when the business cycle is moving into a contraction and therefore the asset is a business cycle hedge

that public utility stocks serve as a hedging asset as they are not viewed as defensive stocks (they do not rise in value during downturns in the stock market) due to asymmetric regulation and returns as discussed in detail in Kolbe and Tye (1990). Under asymmetric regulation, utility regulators have a tendency to allow the return on equity to fall below the allowed return during downturns in the business cycle and to reduce the return should it rise above the allowed return during expansions. Therefore we expect that the parameter estimates of the return-risk relationship to be positive as utility stocks are hypothesized to not be hedges.

We use the GARCH model to estimate the general asset pricing model since the GARCH model accommodates ARCH effects that improve the efficiency of the parameter estimates. It also provides a volatility forecasting model for the conditional volatility of the asset's risk premium. The conditional volatility projection is used, in turn to predict the expected risk premium. We also use the GARCH-in-Mean model (GARCH-M) since it specifies that the conditional expected risk premium is a linear function of its conditional volatility. There is a vast body of literature that estimates asset pricing models with the GARCH and GARCH-M methods and therefore we will not attempt to summarize them here.

The GARCH-M model was initially developed and tested by Engle et al (1987) to estimate the relationship between US Treasury and corporate bond risk premia and their expected volatilities. The GARCH-M model is specified as:

$$R_{t+1} - R_{f\,t+1} = \alpha\sigma_{t+1}^2 + \varepsilon_{t+1} \tag{2}$$

$$\sigma_{t+1}^2 = \beta_0 + \beta_1\sigma_t^2 + \beta_2\varepsilon_t^2 + \eta_{t+1} \tag{3}$$

$$\varepsilon_t\,|\psi_{t-1} \sim T(0, \sigma_t^2) \tag{4}$$

where R_{t+1} is the expected total return on the public utility stock index or individual utility stock; $R_{f\,t+1}$ is the risk-free rate of return or the yield on an index of public utility bonds of a specified bond rating for the equity-to-debt premium; σ_{t+1}^2 is the conditional or predicted variance of the risk premium that is conditioned on past information (ψ_{t-1}), and ε_t is the error term that is conditional on ψ_{t-1}

The conditional distribution of the error term is specified as the non-unitary variance T-distribution due to the thick-tailed distribution of the risk premia data. If the error distribution is thick-tailed, using an approximating distribution that accommodates thick tails improves the efficiency of the estimates. The parameter, α, is the return-to-risk coefficient as specified in Eq 1 as:

$$\alpha = -\frac{vol_t[M_{t+1}]}{E_t\left[M_{t+1}\right]} corr_t[M_{t+1}, R_{i,t+1}] \tag{5}$$

Note that the coefficient will be positive if the conditional correlation between the SDF and the asset return is negative, indicating that the stock is not a hedging asset. Recall that the SDF is the ratio of intertemporal marginal utilities. Assuming a concave utility function, an upward shock in the ratio implies falling consumption, therefore an associated rise (positive correlation) in the return (R_t) would offset the reduction in consumption, thereby causing the sign of α to be negative. The parameter, α, is also the ratio of risk premium to variance, or, the Sharpe ratio

The intercept in Eq. 2 is restricted to zero as specified by the general asset pricing model specification. The restriction on the intercept equal to zero has been found to be robust in producing consistently positive and significant relationships between equity risk premia and risk in GARCH-M models. This is discussed in Lanne and Saikkonen (2006) and Lanne and Luoto (2007). We have found the same results in our modeling in this paper, although we have excluded these results for brevity (available upon request) Therefore we specify the prior assumption that the intercept or the "excess" return, i.e., the return not associated with risk to be equal to zero and drop the intercept from the model.

The consumption asset pricing model is estimated in the empirical section of the paper and applied in the applications section of the paper. The model is tested to (1) determine if equity-to-debt risk premium indices for utilities of differing risk specified by differing bond ratings are validated by the asset pricing model and therefore have some empirical support for risk premium prediction and application to utility cost of capital estimation, (2) determine whether equity risk premia can be predicted and fit the model and therefore be used to estimate the cost of common equity, (3) empirically test the consumption asset pricing model, and (4) ascertain whether utility stocks are assets that hedge shocks to the marginal utility of consumption.

If utility stocks are hedging assets then the cost of common equity should reflect a downward adjustment to a specified risk-free rate to reflect investors' preferences for a hedge and the compensation that they are willing to pay for it.

3 Data and empirical results

We use portfolios as represented by public utility stock and bond indices to estimate the conditional return-risk relationship for the equity-to-debt premium The equity-to-debt risk premium data employed for estimating Eq. 1 with the GARCH-M conditional return-risk regressions are monthly total returns on the Standard and Poor's Public Utilities Stock Index (utility portfolio), and the monthly Moody's Public Utility Aa, A, and Baa yields for the debt cost. We also obtained equity risk premia for the utility portfolio using the Fama-French specified risk-free rate of return, which is the holding period return on a 1-month US Treasury Bill. The data range from January 1928 to December 2007 with 960 observations. The return-risk relationships for the equity-to-debt premia are risk-differentiated by their own bond rating

As a check, we also estimate Eq. 1 with the GARCH-M for large common stock returns using the monthly Ibbotson Large Company Common Stocks Portfolio total returns and the Ibbotson US Long-Term Government income returns as the risk-free rate Additionally, as another check, we do the same for the University of Chicago's Center for Research in Security Prices value-weighted stock index (CRSP) using the Fama-French risk-free rate This is the Fama-French specification of the market equity risk premium. The data range from January 1926 to December 2007 with 984 observations for the Large Company Common Stock estimation and the data ranges

Table 1 Descriptive statistics: public utility and large company common stocks equity-to-debt and equity risk premia

Utility bond rating	Mean	Std. Dev	Skewness	Kurtosis	JB
Aa	0.0037	0.0568	0.0744	10.07	2,001.2***
A	0.0035	0.0568	0.0632	10.06	1,991.8***
Baa	0.0031	0.0568	0.0375	10.02	1,973.6***
Ibbotson					
Large common stocks	0.0054	0.0554	0.4300	12.84	3,954.7***
CRSP value-weighted stock index	0.0062	0.0544	0.2309	10.92	2,519.1

The public utility equity-to-debt risk premia monthly time series is from January 1928 to December 2007 with 960 observations. The equity risk premium monthly time series for the Large Common Stocks and the CRSP index are January 1926 to December 2007 with 984 observations, and January 1926 to December 2007 with 984 observations, respectively. The public utility stocks equity-to-debt risk premia are calculated as the total return on the S&P Public Utilities Index of stocks minus the Moody's Public Utility Aa, A, and Baa Indices yields to maturity. The Large Company Common Stock equity risk premia are the monthly total returns on the Ibbotson Large Company Common Stocks Portfolio minus the Ibbotson Long-Term US Government Bonds Portfolio income yield. The CRSP equity risk premia, or the Fama-French market risk premia are the CRSP total returns on the value-weighted equity index minus the 1-month holding period return on a 1 month Treasury Bill. The Jarque-Bera (JB) statistic is a goodness-of-fit measure of the departure of the distribution of a data series from normality, based on the levels of skewness and excess kurtosis. The JB statistic is χ^2 distributed with 2° of freedom. *** Significant at 0.01 level, one-tailed test

from January 1928 to January 2007 with 960 observations (same as the utilities) for the CRSP estimation.

Table 1 displays the descriptive statistics for these data. We have estimated the mean, standard deviation, skewness and kurtosis parameters, as well as the Jarque-Bera (JB) statistic to test the distribution of the data. The means of the utility equity-to-debt risk premia fall as the risk (bond rating) declines. This is consistent with the notion that larger yields are subtracted from stock returns the lower the bond rating. Intertemporally, there is an inverse relationship between risk premia and interest rates (See Brigham et al. (1985) and Harris et al. (2003)). The mean for risk premia will have a tendency to be larger during low interest rate periods.

Not surprisingly, large company common stocks have the highest mean risk premia as the majority of these firms are not rate-of-return regulated firms with a ceiling on their ROE's close to their cost of capital. Interestingly, the standard deviations of the utility stock returns are similar and slightly higher than large company common stocks. Skewness coefficients are small and positive except for Ibbotson large company common stock returns and CRSP returns that have large positive skewness. This suggests that large unregulated stocks have a tendency to have more and larger positive shocks in returns than do utilities that are rate of return regulated. The kurtosis values show that all of the risk premia are thick-tail distributed. This is also found in the significant JB statistics that test the null hypothesis that the data are normally distributed. The null hypothesis is rejected for all assets. The high kurtosis, low skewness, and significant JB statistics show that the risk premia data are substantially thick-tailed, except for non-utility stocks that are both skewed and thick-tailed. Therefore, robust estimation methods are required to produce efficient regression estimates with non-normal data. Additionally, although not shown but available upon request, the serial correlation and

ARCH Lagrange Multiplier tests show that residuals from OLS regressions of risk premia on volatilities follow an ARCH process. Therefore, the GARCH-M method will improve the efficiency of the estimates. We specify the regression error distribution as a non-unitary variance T-distribution so that thick-tails could be accommodated in the estimation and therefore produce increasingly efficient parameter estimates.

We used maximum likelihood estimation with the likelihood function specified with the non-unitary-variance T-distribution as the approximating distribution of the residuals to accommodate the thick-tailed nature of the error distribution. The equations are estimated as a system using the Marquardt iterative optimization algorithm. The chosen software for estimating the model was EViews© version 6.0 (2007).

Table 2 shows the GARCH-M estimations for the consumption asset pricing Eq. 1. We have estimated Eq 1 for the utility equity risk premia using the Fama-French risk-free rate in addition to the equity-to-debt risk premia risk-differentiated by bond ratings and the two measures of the market equity risk premium. The chosen measure of volatility is the variance of risk premium (in contrast to other such measures such as the standard deviation or the log of variance. Although these results are not shown for brevity, they are robust to these other measures of volatility). The slope, which is the predicted return-to-predicted risk coefficient and Sharpe ratio, is positive and significant at the 99% level for all assets except the utility stock returns with Baa bonds, which is significant at the 95% level. Given that all slopes are positive, public utility stocks are not found to hedge shocks to the marginal utility of consumption. Note that the reward-to-risk slope rises as bond rating rises. This suggests that lower risk utility stocks provide a higher incremental risk-premium for an increase in conditional volatility. This is consistent with other studies that find that lower risk assets, such as shorter maturity bonds, have higher Sharpe Ratios than long-term bonds and stocks. See Pilotte and Sterbenz (2006) and Michelfelder and Pilotte (2011)

The variance equation shows that all GARCH coefficients (β's) are significant at the 1% level and the sums of β_1 and β_2 are close to, but less than 1.0, indicating that the residuals of the risk premium equation follow a GARCH process and that the persistence of a volatility shock on returns and stock prices for utility stocks is temporary. The estimates of the non-unitary variance T-distribution degrees of freedom parameter are low and statistically significant, indicating that the residuals are well approximated by the T. Similar values for the log-likelihood functions (Log-L) show that each of the regressions has a similar goodness-of-fit. Chi-squared distributed likelihood ratio tests (not shown but available upon request) that compare the goodness of fit among the T and normal specifications of the likelihood function of the GARCH-M regressions show that the T has a significantly better fit than the normal distribution.

The GARCH-M results for the large company common stocks portfolio are similar to those of the utility stocks Not surprisingly, large company common stocks do not hedge shocks to the marginal utility of consumption and volatility shocks temporarily affect their valuations The exception is that the return-risk slope is substantially higher than utility stock slopes. This is partially due to the risk-free nature of the risk-free rates used with the non-utility equity risk premia compared to the

Table 2 Estimation of return-risk relation: public utility and large company common stocks

Utility bond rating	α	β_0	β_1	β_2	Log-L	T dist. D.F.
Aa	1.5183***	0.0000***	0.8791***	0.1031***	1,604.4	9.9254***
	(0.5308)	(0.0000)	(0.0230)	(0.0219)		(3.0272)
A	1.4536***	0.0000***	0.8790***	0.1033***	1,605.0	9.9381
	(0.5308)	(0.0000)	(0.0230)	(0.0220)		(3.0408)
Baa	1.3318***	0.0000***	0.8789***	0.1040***	1,605.2	10.0***
	(0.5303)	(0.0000)	(0.0229)	(0.0220)		(3.0540)
Fama-French R_f	2.1428***	0.0000	0.8811***	0.0979***	1,601.0	9.8773***
	(0.5318)	(0.0000)	(0.0232)	(0.0212)		(2.9700)
Ibbotson						
Large company common stocks	2.7753***	0.0001***	0.8381***	0.1186***	1,620.8	8.8457***
	(0.5513)	(0.0000)	(0.0269)	(0.0332)		(2.1613)
CRSP value-weighted stock index	3.3873***	0.0001***	0.8330***	0.1149***	1,598.9	8.8571***
	(0.5673)	(0.0000)	(0.0270)	(0.0358)		(1.9505)

The results below are the GARCH-in-Mean regressions for the risk premium ($R_{t+1} - R_{f\,t+1}$) on the conditional variance of the risk premium (σ^2_{t+1}) in the mean equation. The intercept in the mean equation is restricted to be equal to zero. The public utility equity-to-debt risk premia monthly time series is from January 1928 to December 2007 with 960 observations. The equity risk premium monthly time series for the Large Company Common Stocks and the CRSP index are January 1926 to December 2007 with 984 observations, and January 1926 to December 2007 with 984 observations respectively. The public utility stocks equity-to-debt risk premia are calculated as the total return on the S&P Public Utilities Index of stocks minus the Moody's Public Utility Aa, A, and Baa Indices yields to maturity. The Large Company Common Stock equity risk premia are the monthly total returns on the Ibbotson Large Company Common Stocks Portfolio minus the Ibbotson Long-Term US Government Bonds Portfolio income yield. The CRSP equity risk premia, or the Fama-French market risk premia are the CRSP total returns on the value-weighted equity index minus the 1-month holding period return on a 1 month Treasury Bill. The estimated model is

$$R_{t+1} - R_{f\,t+1} = \alpha\sigma^2_{t+1} + \varepsilon_{t+1} \quad \text{where } \alpha = -\frac{vol_t[M_{t+1}]}{E_t[M_{t+1}]}corr_t[M_{t+1}, R_{t+1}]$$

$$\sigma^2_{t+1} = \beta_0 + \beta_1\sigma^2_t + \beta_2\varepsilon^2_t + \eta_{t+1}$$

The conditional distribution of the error term is the non-unitary variance T-distribution to accommodate the kurtosis of the risk premia and error term. Standard errors are in parentheses. ***, **, * denote significance at the 0.01, 0.05, and 0.10 levels, respectively for two-tail tests

utility bond yields that reflect risk. The utility stocks slope value of 2.1428 using the Fama-French risk-free rate is closer to the higher CRSP value of 3.3873 that is also based on the Fama-French risk-free rate. This is inconsistent with previous results herein and in other papers that find that Sharpe Ratios are lower for higher risk assets unless this finding can be interpreted as utility stocks having more risk than non-regulated stocks. The standard deviations on Table 1 suggest that utility stock return volatilities are as high as the stock returns of non-regulated firms. However, similar model estimates of portfolios of common stocks yield unstable results, such as negative as well as positive return-risk slopes when the intercept is not restricted to zero. See Campbell (1987), Glosten et al. (1993), Harvey (2001), and Whitelaw (1994).

Stock market results are highly sensitive to empirical model specification. Many studies do not consider the impact of a zero-intercept prior restriction on the stability of their results. This simple innovation has led to more consistent results in modeling stock market risk-return relationships, and therefore we have included it in this paper.

The estimation of the consumption asset pricing model for utility stock equity-debt risk premia shows that the use of bond-rating risk-differentiated risk premia are validated as their risk-return relationships are well-fitted by theoretical and empirical models of risk and return. Therefore, these data impound good representations of the risk and reward relationship.

One concern is the intertemporal stability of the alphas Figure 1 plots the utility stock portfolio alpha (using the Fama-French R_f to calculate the premium) and its standard error for 240 month rolling regressions of the model estimated with GARCH-M in the same manner as described above to review the intertemporal stability of the alpha. A 20-year period was used for each estimation to trade off timeliness with sufficient observation of up and down stock market regimes and business cycles. This resulted in 720 estimated alphas from 1947 to 2007. The results show that the utility alpha is stable to the extent that the algebraic sign is always positive and generally significant, therefore the nature of utility stocks are assets that are not and have never been hedges during the second half of the twentieth century up to the present. The value of the alpha does change substantially. The mean of the alpha is 4.40 with a range from -0.11 (insignificantly different from 0) to 11.66. As a comparison, the alpha for the CRSP value-weighted stock index was also estimated with rolling regressions in the same manner and for the same time period. Figure 2 is a plot of the CRSP alpha and standard error. Note that the general stock market alpha is similar to that of utility stocks. They are all positive and almost all statistically significant and follow a strikingly similar cycle. Figure 3 plots both the utility and stock market alphas and demonstrates the similarity. The correlation coefficient between the utility and stock market alphas is 0.88. Recalling that the alpha is a Sharpe Ratio, we see that return to risk ratio does change substantially. This is consistent with the results in Pilotte and Sterbenz (2006).

One other interesting observation is that the standard errors of the alphas are highly stable over the study period and are very similar in magnitude regardless of the size of the corresponding alpha Whereas the alpha follows a cyclical pattern, the volatility in alpha is highly stationary around a constant, long-run mean.

The GARCH-M model estimations of the consumption asset pricing model were specified with variance as the measure of volatility. We also performed the same model estimations with alternative specifications of volatility such as the standard deviation and the log of variance and the results were not sensitive to this specification.

4 Application

We apply the model in this section to compare the cost of common equity capital estimates with the DCF and CAPM models. Using EViews© Version 6.0, we estimated the model coefficients (α, β's) over rolling 24 month periods ending December 2008.

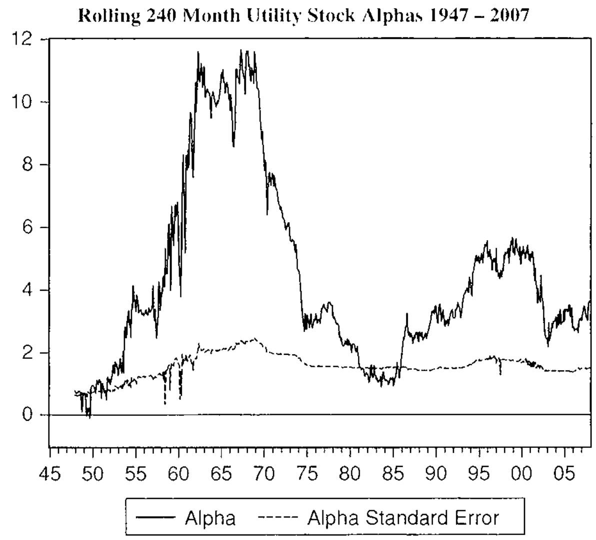


Fig. 1 Rolling 240 month utility stock alphas 1947–2007

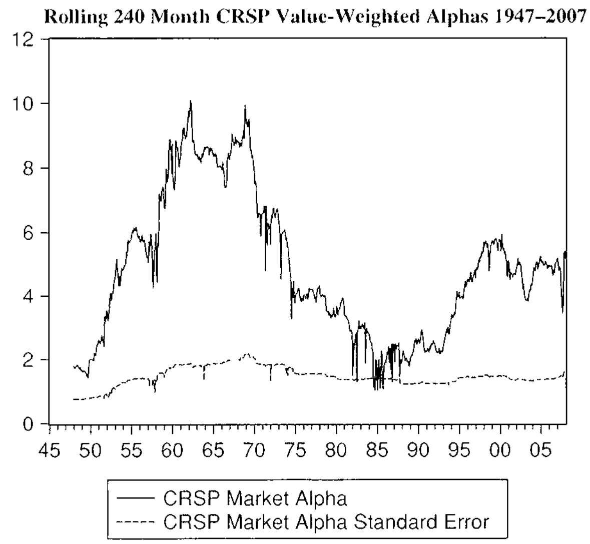


Fig. 2 Rolling 240 month CRSP value-weighted alphas 1947–2007

We repeated the estimation over 5, 10, 15, 20 and 79 year periods.[3] Predicted monthly variances (σ^2_{t+1}) were generated from these estimations to produce predicted risk premiums that were calculated by multiplying the predicted variance by the "α" slope

[3] We did not include the results of the 10 and 15 year estimations to abbreviate the amount of empirical results presented since they added no material insights beyond those already presented

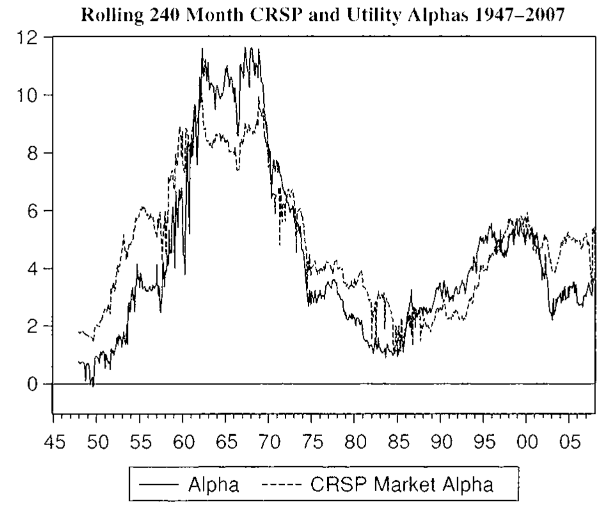


Fig. 3 Rolling 240 month CRSP and utility alphas 1947–2007

Table 3 Estimates of expected risk premia

	Mean (%)		Range (%)		Standard deviation (%)	
	Average	Spot	Average	Spot	Average	Spot
Ibbotson Associates data						
79-years	9 59	5 76	8 74–9 96	2 62–22 60	0 32	5 24
20-years	6 77	6 94	4 99–8 50	2 24–28 95	0 95	6 88
5-years	4 20	10 25	−98 49–11 62	−100 00–39 65	22 00	26 61
S&P Utility Index						
79-years	5 28	2 90	4 30–5 28	1 65–8 15	0 32	1 60
20-years	3 93	3 51	2 78–5 03	2 18–6 88	0 57	1 11
5-years	31 82	326 63	7 77–156 97	6 12–6465 74	31 47	1283 51

coefficient To test the stability of the predicted risk premia over time, the predicted risk premia were calculated using either the predicted variance over each entire time period or the last monthly (spot) predicted variance. Table 3 presents the mean predicted risk premia, the range of predicted premia and the standard deviations for each time period. It is clear from the results that the risk premia are more stable over the rolling 24 month period when calculated using the average predicted variance compared with using the spot variance. Secondly, the 20 and 79 year means are substantially more stable and reasonable in magnitude than the 5 year means.

Next, given the lessons from the analyses above, we apply the model to mechanically[4] estimate the cost of common equity for 8 utility companies using the model and

[4] The term "mechanically" in this context means that the resulting values have been developed in a consistent manner with the same inputs across all utility stocks but no subjective judgment was used to develop final values for each specific utility stock application

the DCF and CAPM as comparisons. We also calculated the realized market return for comparison. Two publicly-traded electric, electric and gas combination, gas, and water utilities respectively were chosen for the application. The Gordon (1974) DCF and CAPM models are used in many utility regulatory jurisdictions in the US.

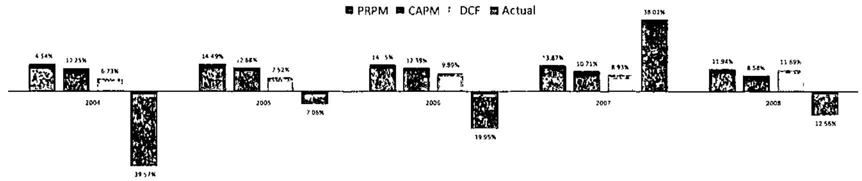
The DCF was applied using a dividend yield, D_0/P_0, derived by dividing the year-end indicated dividend per share (D_0) by the year-end spot market price (P_0). The dividend yield is grown by the year-end I/B/E/S five year projected earnings per share growth rate (g) to derive $D_0(1+g)/P_0$. The one-year predicted dividend yield is then added to the I/B/E/S five-year projected EPS growth rate to obtain the DCF estimate of the cost of common equity capital, k. This study was conducted for the 5 years ending 2008.

The CAPM was applied by multiplying the Value Line beta (β) available at year-end for each company by the long-term historic arithmetic mean market risk premium ($R_m - R_f$). $R_m - R_f$ is derived as the spread of the total return of large company common stocks over the income return on long-term government bonds from the Ibbotson SBBI 2009 Valuation Yearbook The resulting company-specific market equity risk premium is then added to a projected consensus estimate of the yield on 30-year U.S Treasury rate provided by Blue Chip Financial Forecasts as the risk-free rate (R_f) to obtain the CAPM result This study was also conducted over the 5 years ending 2008.

Figures 4–11 show the histograms of the cost of common equity capital estimations for each of the eight public utility stocks and the realized market returns in the forthcoming year. The consumption asset pricing model appears to track more consistently with the CAPM than with the DCF which seems to produce generally lower values than the other methods. The consumption asset pricing model results are similar to the CAPM. The model and the CAPM compete as the best predictor of the rate of return on the book value of common equity (not shown but available upon request), but none of the expected returns were good predictors of market returns. That does not infer that they were not good predictors of *expected* market returns. These results are an initial indicator that the consumption asset pricing model provides reasonable and stable results. This paper does not suggest at this early juncture that the consumption asset pricing model is superior to the CAPM or DCF, although it is based on far less restrictive assumptions than these other models. For example, both the DCF and CAPM assume that markets are efficient. Many assume that the DCF requires that the market-to-book ratio to always equal one, whereas the long-term value for the Standard and Poor's 500 is equal to 2.34. The CAPM assumes that investors demand higher returns for higher volatility and that the minimum required return is the risk-free rate, whereas the consumption asset pricing model allows for investors to require returns less than the risk-free rate for stocks that may have relatively higher volatility but are hedging assets that have desirable return fluctuation patterns that offset downturns in the business cycle. Unlike the CAPM, the model prices the risk to which investors are actually exposed, whether it's systematic risk or not. Some investors are diversified and some are not; the model prices whatever risk to which the aggregate of investors of the specific stock is exposed.

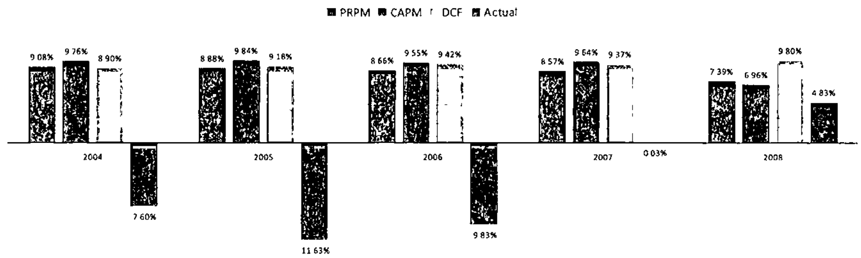
We find that the consumption asset pricing model should be used in combination with other cost of common equity pricing models as additional information in the devel-

Cost of Common Equity Results for Edison International Compared to Market Return*



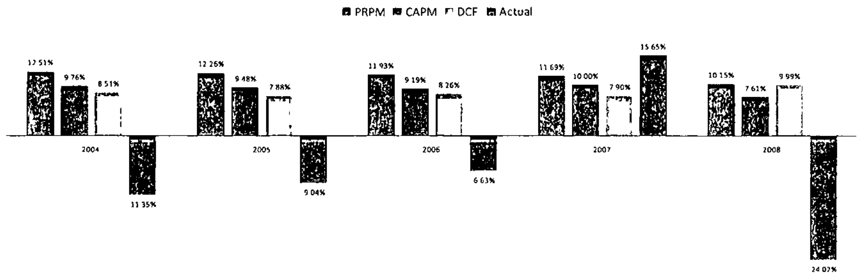
* Market returns calculated for the following years 2005 -2009

Cost of Common Equity Results for Southern Company Compared to Market Return*



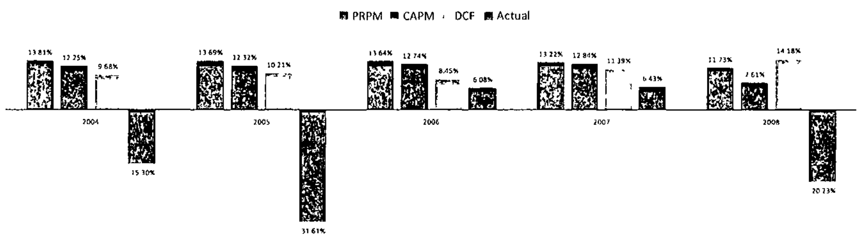
* Market returnscalculated for the following years 2005 -2009

Cost of Common Equity Results for Consolidated Edison Compared to Market Return*



* Market returns calculated for the following years 2005 - 2009

Cost of Common Equity Results for PG&E Corp Compared to Market Return*

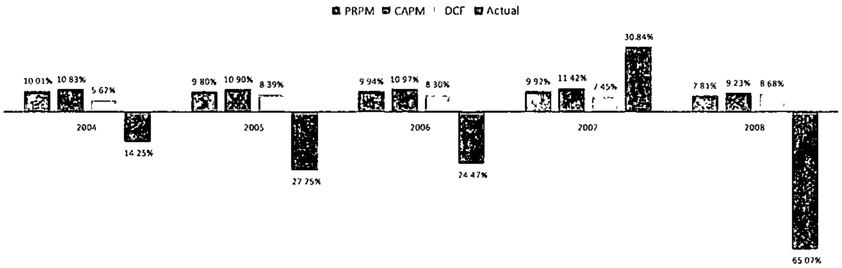


* Market returns calculated for the following years 2005 -2009

Figs. 4–11 Comparison of the cost of common equity estimates and market

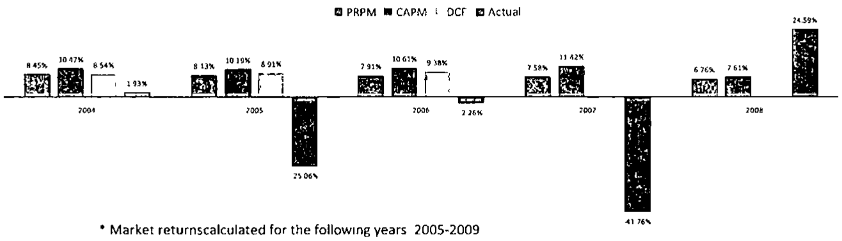
opment of a cost of common equity capital recommendation. Practitioners may find the modeling methods and the use of relatively advanced econometric methods rather cumbersome. The software for performing these estimations is readily available from EViews© and SAS©; two commonly available software packages at utilities, consult-

Cost of Common Equity Results for National Fuel Gas Co. Compared to Market Return*



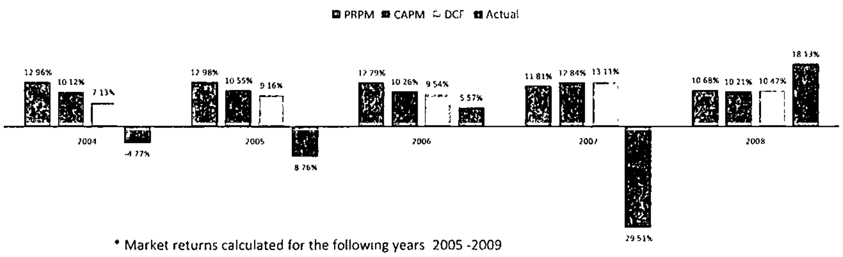
* Market returrnscalculated for the following years. 2005 -2009

Cost of Common Equity Results for Laclede Group Compared to Market Return*



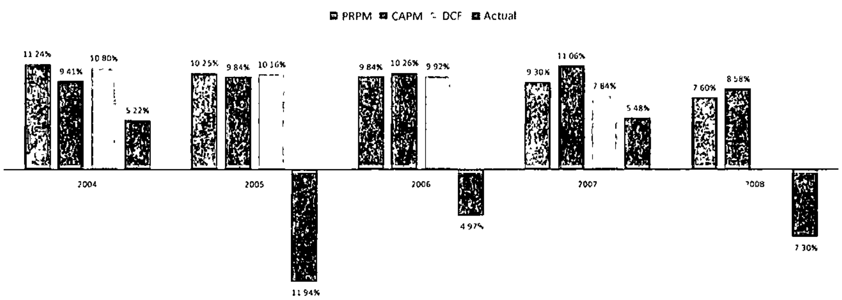
* Market returnscalculated for the following years 2005-2009
Missing DCF Cost of Capital Estimates Due to Unavailable Growth Rate

Cost of Common Equity Results for California Water Service Group Compared to Market Return *



* Market returns calculated for the following years 2005 -2009

Cost of Common Equity Results for Middlesex Water Company Compared to Market Return *



* Market returnscalculated for following years 2005 -2009
Missing DCF Cost of Capital Estimate Due to Unavailable Growth Rate

Figs. 4–11 continued

ing firms and financial firms. Recent Ph.D. and M.S. holding members of research departments of investment and consulting firms have ready access to the model and methods discussed in this paper, although it will require years for these tools, like any "new" technology, to diffuse into standard use. Another problem is that the model requires a substantial time series history on stock returns data to develop stable estimates of risk premia This is problematic especially for the electric and gas utility industries that have consolidated with many mergers in the recent past. This problem can be addressed by developing and predicting the value-weighted risk premium of a portfolio of similar stocks such as electric utilities that have nuclear generating assets. The specific stock in question would be included in the returns index with a weight based on market capitalization that would go to 0 when the stock price history is no longer existent reaching back into the past.

5 Conclusion

The purpose of this paper is to introduce, test empirically and apply a general consumption based asset pricing model that is based on a minimum of assumptions and restrictions that can be used to predict the risk premium to be applied in estimating the cost of common equity for public utilities in regulatory proceedings The results support the simple consumption-based asset pricing model that predicts the ex ante risk premium with a conditionally predicted volatility in risk premium. The estimates of the cost of common equity from the consumption asset pricing model compare well with rates of return on the book value of common equity and with the CAPM, although both the model and the CAPM results are substantially higher than the DCF. This is quite common in the practice of the cost of common equity in the utility industry. The results of the model are stable and consistent over time. Therefore the model should be considered as it provides additional evidence on the cost of common equity in general and specifically in public utility regulatory proceedings. Secondly, the use of bond-rated yields to predict risk differentiated equity-to-debt risk premia is supported by the empirical evidence and therefore should be applied in estimating the cost of common equity. Finally, the robust empirical evidence on the positive risk-return relationship also shows that utility stocks are not a consumption hedge and are not good hedging securities against contractions in the economy The model and estimation methodology presented in this paper provide a relatively simple tool to determine whether any asset is a hedge to adverse changes in the business cycle through the level of consumption in the economy.

Acknowledgments We would like to thank Dylan D'Ascendis Sal Giunta Selby Jones, III and Alison McVicker for highly capable research assistance, participants at the Center for Research in Regulated Industries Eastern Conferences and the Society of Utility Regulatory and Financial Analysts Annual Financial Forum, two anonymous reviewers and the editor for helpful comments

References

Brigham, E F, Shome D K & Vinson, S R (1985) The risk premium approach to measuring a utility's cost of capital *Financial Management 14* 33–45

Campbell J Y (1987) Stock returns and the term structure *Journal of Financial Economics, 18* 373–399
Campbell, J Y (1993) Intertemporal asset pricing without consumption data *American Economic Review, 83*, 487–512
Campbell, J Y, & Cochrane J H (1999) By force of habit A consumption-based explanation of aggregate stock market behavior *Journal of Political Economy 107*, 205–251
Cochrane, J H (2004) *Asset pricing* Revised Edition Princeton, NJ Princeton University Press
Cochrane, J H (2006) Financial markets and the real economy NBER Working Paper
Cochrane J H (2007) *Portfolio theory Manuscript* University of Chicago
Engle, R F Lilein, D, & Robins, R (1987) Estimation of time varying risk premia in the term structure The ARCH-M model *Econometrica, 55*, 391–407
EViews© (2007) Version 6 0 Quantitative Micro Software, LLC
Fama, E, & French, K (2004) The capital asset pricing model Theory and evidence *Journal of Economic Perspectives, 18*, 25–46
Glosten L R, Jaganathan, R, & Runkle, D E (1993) Relationship between the expected value and the volatility of the nominal excess returns on stocks *Journal of Finance, 48*, 1779–1801
Gordon, M, (1974) *The cost of capital to a public utility* East Lansing, MI MSU Public Utility Studies
Harris, R S, Marston, F C, Mishra, D R, & O'Brien T J (2003) Ex ante cost of equity estimate of S&P 500 firms The choice between global and domestic CAPM *Financial Management, 32*, 51–66
Harvey, C R (2001) The specification of conditional expectations *Journal of Empirical Finance, 8* 573–637
Kolbe, A L & Tye W B (1990) The *Duquense* opinion How much "Hope" is there for investors in regulated firms *Yale Journal on Regulation, 8*, 113–157
Lanne, M, & Luoto, J (2007) Robustness of risk-return relationship in the U S stock market Helsinki Center of Economic Research, Discussion Paper No 168
Lanne M, & Saikkonen, P (2006) Why is it so difficult to uncover the risk-return tradeoff in stock returns? *Economic Letters, 92*, 118–125
Merton, R C (1973) An intertemporal capital asset pricing model *Econometrica, 41*, 867–887
Michelfelder, R A, & Pilotte E A (2011) Treasury bond risk and return, the implications for the hedging of consumption and lessons for asset pricing *Journal of Economics and Business* (forthcoming)
Pilotte, E, & Sterbenz, F (2006) Sharpe and treynor ratios on treasury bonds *Journal of Business, 79*, 149–180
Whitelaw, R W (1994) Time-variation and covariations in the expectation and volatility of stock market returns *Journal of Finance, 49*, 515–541

Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation

Robert F. Engle

Econometrica, Volume 50, Issue 4 (Jul., 1982), 987-1008.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198207%2950%3A4%3C987%3AACHWEO%3E2.0.CO%3B2-3

Econometrica, Vol. 50, No. 4 (July, 1982)

AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY WITH ESTIMATES OF THE VARIANCE OF UNITED KINGDOM INFLATION[1]

BY ROBERT F. ENGLE

Traditional econometric models assume a constant one-period forecast variance. To generalize this implausible assumption, a new class of stochastic processes called autoregressive conditional heteroscedastic (ARCH) processes are introduced in this paper. These are mean zero, serially uncorrelated processes with nonconstant variances conditional on the past, but constant unconditional variances. For such processes, the recent past gives information about the one-period forecast variance.

A regression model is then introduced with disturbances following an ARCH process. Maximum likelihood estimators are described and a simple scoring iteration formulated. Ordinary least squares maintains its optimality properties in this set-up, but maximum likelihood is more efficient. The relative efficiency is calculated and can be infinite. To test whether the disturbances follow an ARCH process, the Lagrange multiplier procedure is employed. The test is based simply on the autocorrelation of the squared OLS residuals.

This model is used to estimate the means and variances of inflation in the U.K. The ARCH effect is found to be significant and the estimated variances increase substantially during the chaotic seventies.

1. INTRODUCTION

IF A RANDOM VARIABLE y_t is drawn from the conditional density function $f(y_t \mid y_{t-1})$, the forecast of today's value based upon the past information, under standard assumptions, is simply $E(y_t \mid y_{t-1})$, which depends upon the value of the conditioning variable y_{t-1}. The variance of this one-period forecast is given by $V(y_t \mid y_{t-1})$. Such an expression recognizes that the conditional forecast variance depends upon past information and may therefore be a random variable. For conventional econometric models, however, the conditional variance does not depend upon y_{t-1}. This paper will propose a class of models where the variance does depend upon the past and will argue for their usefulness in economics. Estimation methods, tests for the presence of such models, and an empirical example will be presented.

Consider initially the first-order autoregression

$$y_t = \gamma y_{t-1} + \epsilon_t$$

where ϵ is white noise with $V(\epsilon) = \sigma^2$. The conditional mean of y_t is γy_{t-1} while the unconditional mean is zero. Clearly, the vast improvement in forecasts due to time-series models stems from the use of the conditional mean. The conditional

[1]This paper was written while the author was visiting the London School of Economics. He benefited greatly from many stimulating conversations with David Hendry and helpful suggestions by Denis Sargan and Andrew Harvey. Special thanks are due Frank Srba who carried out the computations. Further insightful comments are due to Clive Granger, Tom Rothenberg, Edmond Malinvaud, Jean-Francois Richard, Wayne Fuller, and two anonymous referees. The research was supported by NSF SOC 78-09476 and The International Centre for Economics and Related Disciplines. All errors remain the author's responsibility.

variance of y_t is σ^2 while the unconditional variance is $\sigma^2/1-\gamma^2$. For real processes one might expect better forecast intervals if additional information from the past were allowed to affect the forecast variance; a more general class of models seems desirable.

The standard approach of heteroscedasticity is to introduce an exogenous variable x_t which predicts the variance. With a known zero mean, the model might be

$$y_t = \epsilon_t x_{t-1}$$

where again $V(\epsilon) = \sigma^2$. The variance of y_t is simply $\sigma^2 x_{t-1}^2$ and, therefore, the forecast interval depends upon the evolution of an exogenous variable. This standard solution to the problem seems unsatisfactory, as it requires a specification of the causes of the changing variance, rather than recognizing that both conditional means and variances may jointly evolve over time. Perhaps because of this difficulty, heteroscedasticity corrections are rarely considered in time-series data.

A model which allows the conditional variance to depend on the past realization of the series is the bilinear model described by Granger and Andersen [**13**]. A simple case is

$$y_t = \epsilon_t y_{t-1}.$$

The conditional variance is now $\sigma^2 y_{t-1}^2$. However, the unconditional variance is either zero or infinity, which makes this an unattractive formulation, although slight generalizations avoid this problem.

A preferable model is

$$y_t = \epsilon_t h_t^{1/2},$$

$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2,$$

with $V(\epsilon_t) = 1$. This is an example of what will be called an autoregressive conditional heteroscedasticity (ARCH) model. It is not exactly a bilinear model, but is very close to one. Adding the assumption of normality, it can be more directly expressed in terms of ψ_t, the information set available at time t. Using conditional densities,

$$y_t \mid \psi_{t-1} \sim N(0, h_t), \tag{1}$$

$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2. \tag{2}$$

The variance function can be expressed more generally as

$$h_t = h(y_{t-1}, y_{t-2}, \ldots, y_{t-p}, \alpha) \tag{3}$$

where p is the order of the ARCH process and α is a vector of unknown parameters.

The ARCH regression model is obtained by assuming that the mean of y_t is given as $x_t\beta$, a linear combination of lagged endogenous and exogenous variables included in the information set ψ_{t-1} with β a vector of unknown parameters. Formally,

$$
\begin{aligned}
& y_t \mid \psi_{t-1} \sim N(x_t\beta, h_t), \\
(4) \qquad & h_t = h(\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-p}, \alpha), \\
& \epsilon_t = y_t - x_t\beta.
\end{aligned}
$$

The variance function can be further generalized to include current and lagged x's as these also enter the information set. The h function then becomes

$$(5) \qquad h_t = h(\epsilon_{t-1}, \ldots, \epsilon_{t-p}, x_t, x_{t-1}, \ldots, x_{t-p}, \alpha)$$

or simply

$$h_t = h(\psi_{t-1}, \alpha).$$

This generalization will not be treated in this paper, but represents a simple extension of the results. In particular, if the h function factors into

$$h_t = h_\epsilon(\epsilon_{t-1}, \ldots, \epsilon_{t-p}, \alpha)h_x(x_t, \ldots, x_{t-p}),$$

the two types of heteroscedasticity can be dealt with sequentially by first correcting for the x component and then fitting the ARCH model on the transformed data.

The ARCH regression model in (4) has a variety of characteristics which make it attractive for econometric applications. Econometric forecasters have found that their ability to predict the future varies from one period to another. McNees [17, p. 52] suggests that, "the inherent uncertainty or randomness associated with different forecast periods seems to vary widely over time." He also documents that, "large and small errors tend to cluster together (in contiguous time periods)." This analysis immediately suggests the usefulness of the ARCH model where the underlying forecast variance may change over time and is predicted by past forecast errors. The results presented by McNees also show some serial correlation during the episodes of large variance.

A second example is found in monetary theory and the theory of finance. By the simplest assumptions, portfolios of financial assets are held as functions of the expected means and variances of the rates of return. Any shifts in asset demand must be associated with changes in expected means and variances of the rates of return. If the mean is assumed to follow a standard regression or time-series model, the variance is immediately constrained to be constant over time. The use of an exogenous variable to explain changes in variance is usually not appropriate.

A third interpretation is that the ARCH regression model is an approximation to a more complex regression which has non-ARCH disturbances. The ARCH specification might then be picking up the effect of variables omitted from the estimated model. The existence of an ARCH effect would be interpreted as evidence of misspecification, either by omitted variables or through structural change. If this is the case, ARCH may be a better approximation to reality than making standard assumptions about the disturbances, but trying to find the omitted variable or determine the nature of the structural change would be even better.

Empirical work using time-series data frequently adopts *ad hoc* methods to measure (and allow) shifts in the variance over time. For example, Klein [**15**] obtains estimates of variance by constructing the five-period moving variance about the ten-period moving mean of annual inflation rates. Others, such as Khan [**14**], resort to the notion of "variability" rather than variance, and use the absolute value of the first difference of the inflation rate. Engle [**10**] compares these with the ARCH estimates for U.S. data.

2. THE LIKELIHOOD FUNCTION

Suppose y_t is generated by an ARCH process described in equations (1) and (3). The properties of this process can easily be determined by repeated application of the relation $Ex = E(Ex \mid \psi))$. The mean of y_t is zero and all autocovariances are zero. The unconditional variance is given by $\sigma_t^2 = Ey_t^2 = Eh_t$. For many functions h and values of α, the variance is independent of t. Under such conditions, y_t is covariance stationary; a set of sufficient conditions for this is derived below.

Although the process defined by (1) and (3) has all observations conditionally normally distributed, the vector of y is not jointly normally distributed. The joint density is the product of all the conditional densities and, therefore, the log likelihood is the sum of the conditional normal log likelihoods corresponding to (1) and (3). Let l be the average log likelihood and l_t be the log likelihood of the tth observation and T the sample size. Then

$$\begin{aligned} l &= \frac{1}{T}\sum_{t=1}^{T} l_t, \\ l_t &= -\tfrac{1}{2}\log h_t - \tfrac{1}{2} y_t^2/h_t, \end{aligned} \tag{6}$$

apart from some constants in the likelihood.

To estimate the unknown parameters α, this likelihood function can be maximized. The first-order conditions are

$$\frac{\partial l_t}{\partial \alpha} = \frac{1}{2h_t}\frac{\partial h_t}{\partial \alpha}\left(\frac{y_t^2}{h_t} - 1\right) \tag{7}$$

and the Hessian is

$$\frac{\partial^2 l_t}{\partial\alpha\partial\alpha'} = -\frac{1}{2h_t^2}\frac{\partial h_t}{\partial\alpha}\frac{\partial h_t}{\partial\alpha'}\left(\frac{y_t^2}{h_t}\right) + \left[\frac{y_t^2}{h_t} - 1\right]\frac{\partial}{\partial\alpha'}\left[\frac{1}{2h_t}\frac{\partial h_t}{\partial\alpha}\right]. \tag{8}$$

The conditional expectation of the second term, given ψ_{t-m-1}, is zero, and of the last factor in the first, is just one. Hence, the information matrix, which is simply the negative expectation of the Hessian averaged over all observations, becomes

$$\mathcal{I}_{\alpha\alpha} = \sum_t \frac{1}{2T} E\left[\frac{1}{h_t^2}\frac{\partial h_t}{\partial\alpha'}\frac{\partial h_t}{\partial\alpha'}\right] \tag{9}$$

which is consistently estimated by

$$\hat{\mathcal{I}}_{\alpha\alpha} = \frac{1}{T}\sum_t\left[\frac{1}{2h_t^2}\frac{\partial h_t}{\partial\alpha}\frac{\partial h_t}{\partial\alpha'}\right]. \tag{10}$$

If the h function is pth order linear (in the squares), so that it can be written as

$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \cdots + \alpha_p y_{t-p}^2, \tag{11}$$

then the information matrix and gradient have a particularly simple form. Let $z_t = (1, y_{t-1}^2, \ldots, y_{t-p}^2)$ and $\alpha' = (\alpha_0, \alpha_1, \cdots, \alpha_p)$ so that (11) can be rewritten as

$$h_t = z_t\alpha. \tag{12}$$

The gradient then becomes simply

$$\frac{\partial l}{\partial\alpha} = \frac{1}{2h_t} z_t\left(\frac{y_t^2}{h_t} - 1\right) \tag{13}$$

and the estimate of the information matrix

$$\hat{\mathcal{I}}_{\alpha\alpha} = \frac{1}{2T}\sum_t (z_t' z_t / h_t^2). \tag{14}$$

3. DISTRIBUTION OF THE FIRST-ORDER LINEAR ARCH PROCESS

The simplest and often very useful ARCH model is the first-order linear model given by (1) and (2). A large observation for y will lead to a large variance for the next period's distribution, but the memory is confined to one period. If $\alpha_1 = 0$, of course y will be Gaussian white noise and if it is a positive number, successive observations will be dependent through higher-order moments. As shown below, if α_1 is too large, the variance of the process will be infinite.

To determine the conditions for the process to be stationary and to find the marginal distribution of the y's, a recursive argument is required. The odd moments are immediately seen to be zero by symmetry and the even moments are computed using the following theorem. In all cases it is assumed that the process begins indefinitely far in the past with $2r$ finite initial moments.

THEOREM 1: *For integer r, the $2r$th moment of a first-order linear ARCH process with $\alpha_0 > 0$, $\alpha_1 \geq 0$, exists if, and only if,*

$$\alpha_1^r \prod_{j=1}^{r} (2j-1) < 1.$$

A constructive expression for the moments is given in the proof.

PROOF: See Appendix.

The theorem is easily used to find the second and fourth moments of a first-order process. Letting $w_t = (y_t^4, y_t^2)'$,

$$E(w_t \mid \psi_{t-1}) = \begin{pmatrix} 3\alpha_0^2 \\ \alpha_0 \end{pmatrix} + \begin{pmatrix} 3\alpha_1^2 & 6\alpha_0\alpha_1 \\ 0 & \alpha_1 \end{pmatrix} w_{t-1}.$$

The condition for the variance to be finite is simply that $\alpha_1 < 1$, while to have a finite fourth moment it is also required that $3\alpha_1^2 < 1$. If these conditions are met, the moments can be computed from (A4) as

$$E(w_t) = \begin{bmatrix} \left[\dfrac{3\alpha_0^2}{(1-\alpha_1)^2} \right] \left[\dfrac{1-\alpha_1^2}{1-3\alpha_1^2} \right] \\ \dfrac{\alpha_0}{1-\alpha_1} \end{bmatrix}. \tag{15}$$

The lower element is the unconditional variance, while the upper product gives the fourth moment. The first expression in square brackets is three times the squared variance. For $\alpha_1 \neq 0$, the second term is strictly greater than one implying a fourth moment greater than that of a normal random variable.

The first-order ARCH process generates data with fatter tails than the normal density. Many statistical procedures have been designed to be robust to large errors, but to the author's knowledge, none of this literature has made use of the fact that temporal clustering of outliers can be used to predict their occurrence and minimize their effects. This is exactly the approach taken by the ARCH model.

4 GENERAL ARCH PROCESSES

The conditions for a first-order linear ARCH process to have a finite variance and, therefore, to be covariance stationary can directly be generalized for pth-order processes.

THEOREM 2: *The pth-order linear ARCH processes, with $\alpha_0 > 0$, $\alpha_1, \ldots, \alpha_p \geq 0$, is covariance stationary if, and only if, the associated characteristic equation has all roots outside the unit circle. The stationary variance is given by $E(y_t^2) = \alpha_0/(1 - \sum_{j=1}^{p} \alpha_j)$.*

PROOF: See Appendix.

Although the pth-order linear model is a convenient specification, it is likely that other formulations of the variance model may be more appropriate for particular applications. Two simple alternatives are the exponential and absolute value forms:

$$h_t = \exp(\alpha_0 + \alpha_1 y_{t-1}^2), \tag{16}$$

$$h_t = \alpha_0 + \alpha_1 |y_{t-1}|. \tag{17}$$

These provide an interesting contrast. The exponential form has the advantage that the variance is positive for all values of alpha, but it is not difficult to show that data generated from such a model have infinite variance for any value of $\alpha_1 \neq 0$. The implications of this deserve further study. The absolute value form requires both parameters to be positive, but can be shown to have finite variance for any parameter values.

In order to find estimation results which are more general than the linear model, general conditions on the variance model will be formulated and shown to be implied for the linear process.

Let ξ_t be a $p \times 1$ random vector drawn from the sample space Ξ, which has elements $\xi_t' = (\xi_{t-1}, \ldots, \xi_{t-p})$. For any ξ_t, let ξ_t^* be identical, except that the mth element has been multiplied by -1, where m lies between 1 and p.

DEFINITION: The ARCH process defined by (1) and (3) is *symmetric* if

(a) $h(\xi_t) = h(\xi_t^*)$ for any m and $\xi_t \epsilon \Xi$,

(b) $\partial h(\xi_t)/\partial \alpha_i = \partial h(\xi_t^*)/\partial \alpha_i$ for any m, i and $\xi_t \epsilon \Xi$,

(c) $\partial h(\xi_t)/\partial \xi_{t-m} = -\partial h(\xi_t^*)/\partial \xi_{t-m}$ for any m and $\xi_t \epsilon \Xi$.

All the functions described have been symmetric. This condition is the main distinction between mean and variance models.

Another characterization of general ARCH models is in terms of regularity conditions.

DEFINITION: The ARCH model defined by (1) and (3) is *regular* if

(a) $\min h(\xi_t) \geq \delta$ for some $\delta > 0$ and $\xi_t \epsilon \Xi$,

(b) $E(|\partial h(\xi_t)/\partial \alpha_i||\partial h(\xi_t)/\partial \xi_{t-m}| \,|\psi_{t-m-1})$ exists for all i, m, t.

The first portion of the definition is very important and easy to check, as it requires the variance always to be positive. This eliminates, for example, the log-log autoregression. The second portion is difficult to check in some cases, yet should generally be true if the process is stationary with bounded derivatives, since conditional expectations are finite if unconditional ones are. Condition (b) is a sufficient condition for the existence of some expectations of the Hessian used in Theorem 4. Presumably weaker conditions could be found.

THEOREM 3: *The pth-order linear ARCH model satisfies the regularity conditions, if $\alpha_0 > 0$ and $\alpha_1, \ldots, \alpha_p \geq 0$.*

PROOF: See Appendix.

In the estimation portion of the paper, a very substantial simplification results if the ARCH process is symmetric and regular.

5. ARCH REGRESSION MODELS

If the ARCH random variables discussed thus far have a non-zero mean, which can be expressed as a linear combination of exogenous and lagged dependent variables, then a regression framework is appropriate, and the model can be written as in (4) or (5). An alternative interpretation for the model is that the disturbances in a linear regression follow an ARCH process.

In the pth-order linear case, the specification and likelihood are given by

$$\begin{aligned} & y_t \mid \psi_{t-1} \sim N(x_t\beta, h_t), \\ & h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \cdots + \alpha_p\epsilon_{t-p}^2, \\ (18) \qquad & \epsilon_t = y_t - x_t\beta, \\ & l = \frac{1}{T}\sum_{t=1}^{T} l_t, \\ & l_t = -\tfrac{1}{2}\log h_t - \tfrac{1}{2}\epsilon_t^2/h_t, \end{aligned}$$

where x_t may include lagged dependent and exogenous variables and an irrelevant constant has been omitted from the likelihood. This likelihood function can be maximized with respect to the unknown parameters α and β. Attractive methods for computing such an estimate and its properties are discussed below.

Under the assumptions in (18), the ordinary least squares estimator of β is still consistent as x and ϵ are uncorrelated through the definition of the regression as a conditional expectation. If the x's can be treated as fixed constants then the least squares standard errors will be correct; however, if there are lagged dependent variables in x_t, the standard errors as conventionally computed will not be consistent, since the squares of the disturbances will be correlated with squares of the x's. This is an extension of White's [18] argument on heteroscedasticity and it suggests that using his alternative form for the covariance matrix would give a consistent estimate of the least-squares standard errors.

If the regressors include no lagged dependent variables and the process is stationary, then letting y and x be the $T \times 1$ and $T \times K$ vector and matrix of dependent and independent variables, respectively,

$$\begin{aligned} &E(y \mid x) = x\beta, \\ (19) \qquad &\operatorname{Var}(y \mid x) = \sigma^2 I, \end{aligned}$$

and the Gauss–Markov assumptions are statisfied. Ordinary least squares is the best linear unbiased estimator for the model in (18) and the variance estimates are unbiased and consistent. However, maximum likelihood is different and consequently asymptotically superior; ordinary least squares does not achieve the Cramer–Rao bound. The maximum-likelihood estimator is nonlinear and is more efficient than OLS by an amount calculated in Section 6.

The maximum likelihood estimator is found by solving the first order conditions. The derivative with respect to β is

$$(20) \qquad \frac{\partial l_t}{\partial \beta} = \frac{\epsilon_t x_t'}{h_t} + \frac{1}{2h_t}\frac{\partial h_t}{\partial \beta}\left(\frac{\epsilon_t^2}{h_t} - 1\right).$$

The first term is the familiar first-order condition for an exogenous heteroscedastic correction; the second term results because h_t is also a function of the β's, as in Amemiya [1]. Substituting the linear variance function gives

$$(21) \qquad \frac{\partial l}{\partial \beta} = \frac{1}{T}\sum\left[\frac{\epsilon_t x_t'}{h_t} - \frac{1}{h_t}\left(\frac{\epsilon_t^2}{h_t} - 1\right)\sum_j \alpha_j \epsilon_{t-j} x_{t-j}'\right],$$

which can be rewritten approximately by collecting terms in x and ϵ as

$$\begin{aligned} (22) \qquad \frac{\partial l}{\partial \beta} &= \frac{1}{T}\sum_t x_t'\epsilon_t\left[h_t^{-1} - \sum_{j=1}^{p} \alpha_j h_{t+j}^{-2}(\epsilon_{t+j}^2 - h_{t+j})\right] \\ &\cong \frac{1}{T}\sum_t x_t'\epsilon_t s_t. \end{aligned}$$

The Hessian is

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \beta \partial \beta'} = &- \frac{x_t' x_t}{h_t} - \frac{1}{2h_t^2}\frac{\partial h_t}{\partial \beta}\frac{\partial h_t}{\partial \beta'}\left(\frac{\epsilon_t^2}{h_t}\right) \\ &- \frac{2\epsilon_t x_t'}{h_t^2}\frac{\partial h_t}{\partial \beta} + \left(\frac{\epsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial \beta'}\left[\frac{1}{2h_t}\frac{\partial h_t}{\partial \beta}\right]. \end{aligned}$$

Taking conditional expectations of the Hessian, the last two terms vanish because h_t is entirely a function of the past. Similarly, ϵ_t^2/h_t becomes one, since it is the only current value in the second term. Notice that these results hold regardless of whether x_t includes lagged-dependent variables. The information matrix is the average over all t of the expected value of the conditional expectation and is, therefore, given by

$$\mathcal{I}_{\beta\beta} = \frac{1}{T}\sum_t E\left[E\left(\frac{\partial^2 l_t}{\partial\beta\,\partial\beta'} \mid \psi_{t-1}\right)\right] = \frac{1}{T}\sum_t E\left[\frac{x_t'x_t}{h_t} + \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\beta}\frac{\partial h_t}{\partial\beta'}\right]. \tag{23}$$

For the pth order linear ARCH regression this is consistently estimated by

$$\hat{\mathcal{I}}_{\beta\beta} = \frac{1}{T}\sum\left[\frac{x_t'x_t}{h_t} + 2\sum_j \alpha_j^2\frac{\epsilon_{t-j}^2}{h_t^2}x_{t-j}'x_{t-j}\right]. \tag{24}$$

By gathering terms in $x_t'x_t$, (24) can be rewritten, except for end effects, as

$$\hat{\mathcal{I}}_{\beta\beta} = \frac{1}{T}\sum_t x_t'x_t\left[h_t^{-1} + 2\epsilon_t^2\sum_{j=1}^{p}\alpha_j^2 h_{t+j}^{-2}\right] \equiv \frac{1}{T}\sum_t x_t'x_t r_t^2. \tag{25}$$

In a similar fashion, the off-diagonal blocks of the information matrix can be expressed as:

$$\mathcal{I}_{\alpha\beta} = \frac{1}{T}\sum_t E\left(\frac{1}{2h_t^2}\frac{\partial h_t}{\partial\alpha}\frac{\partial h_t}{\partial\beta'}\right). \tag{26}$$

The important result to be shown in Theorem 4 below is that this off-diagonal block is zero. The implications are far-reaching in that estimation of α and β can be undertaken separately without asymptotic loss of efficiency and their variances can be calculated separately.

THEOREM 4: *If an ARCH regression model is symmetric and regular, then* $\mathcal{I}_{\alpha\beta} = 0$.

PROOF: See Appendix.

6. ESTIMATION OF THE ARCH REGRESSION MODEL

Because of the block diagonality of the information matrix, the estimation of α and β can be considered separately without loss of asymptotic efficiency.

Furthermore, either can be estimated with full efficiency based only on a consistent estimate of the other. See, for example, Cox and Hinkley [**6**, p. 308]. The procedure recommended here is to initially estimate β by ordinary least squares, and obtain the residuals. From these residuals, an efficient estimate of α can be constructed, and based upon these $\hat{\alpha}$ estimates, efficient estimates of β are found. The iterations are calculated using the scoring algorithm. Each step for a parameter vector ϕ produces estimates ϕ^{i+1} based on ϕ^i according to

$$(27) \qquad \phi^{i+1} = \phi^i + \left[\hat{\mathcal{I}}^i_{\phi\phi}\right]^{-1} \frac{1}{T} \sum_t \frac{\partial l^i_t}{\partial \phi},$$

where $\hat{\mathcal{I}}^i$ and $\partial l^i_t/\partial\phi$ are evaluated at ϕ^i. The advantage of this algorithm is partly that it requires only first derivatives of the likelihood function in this case and partly that it uses the statistical properties of the problem to tailor the algorithm to this application.

For the pth-order linear model, the scoring step for α can be rewritten by substituting (12), (13), and (14) into (27) and interpreting y_t^2 as the residuals e_t^2. The iteration is simply

$$(28) \qquad \alpha^{i+1} = \alpha^i + (\tilde{z}'\tilde{z})^{-1}\tilde{z}'f^i$$

where

$$\tilde{z}_t = \left(1, e^2_{t-1}, \ldots, e^2_{t-p}\right)/h^i_t,$$

$$\tilde{z}' = (\tilde{z}'_1, \ldots, \tilde{z}'_T),$$

$$f^i_t = (e^2_t - h^i_t)/h^i_t,$$

$$f^{i\prime} = (f^i_1, \ldots, f^i_T).$$

In these expressions, e_t is the residual from iteration i, h^i_t is the estimated conditional variance, and α^i is the estimate of the vector of unknown parameters from iteration i. Each step is, therefore, easily constructed from a least-squares regression on transformed variables. The variance-covariance matrix of the parameters is consistently estimated by the inverse of the estimate of the information matrix divided by T, which is simply $2(\tilde{z}'\tilde{z})^{-1}$. This differs slightly from $\hat{\sigma}^2(\tilde{z}'\tilde{z})^{-1}$ computed by the auxiliary regression. Asymptotically, $\hat{\sigma}^2 = 2$, if the distributional assumptions are correct, but it is not clear which formulation is better in practice.

The parameters in α must satisfy some nonnegativity conditions and some stationarity conditions. These could be imposed via penalty functions or the parameters could be estimated and checked for conformity. The latter approach is used here, although a perhaps useful reformulation of the model might employ squares to impose the nonnegativity constraints directly:

$$(29) \qquad h_t = \alpha_0^2 + \alpha_1^2\epsilon^2_{t-1} + \cdots + \alpha_p^2\epsilon^2_{t-p}.$$

Convergence for such an iteration can be formulated in many ways. Following Belsley [3], a simple criterion is the gradient around the inverse Hessian. For a parameter vector, ϕ, this is

$$(30)\qquad \theta = \frac{\partial l'}{\partial \phi}\left(\frac{\partial^2 l}{\partial \phi \partial \phi'}\right)^{-1}\frac{\partial l}{\partial \phi}.$$

Using θ as the convergence criterion is attractive, as it provides a natural normalization and as it is interpretable as the remainder term in a Taylor-series expansion about the estimated maximum. In any case, substituting the gradient and estimated information matrix in (30), $\theta = R^2$ of the auxiliary regression.

For a given estimate of α, a scoring step can be computed to improve the estimate of beta. The scoring algorithm for β is

$$(31)\qquad \beta^{i+1} = \beta^i + \left[\hat{\mathcal{I}}_{\beta\beta}\right]^{-1}\frac{\partial l^i}{\partial \beta}.$$

Defining $\tilde{x}_t = x_t r_t$ and $\tilde{e}_t = e_t s_t / r_t$ with $\tilde{x}$ and $\tilde{e}$ as the corresponding matrix and vector, (31) can be rewritten using (22) and (24) and e_t for the estimate of ϵ_t on the ith iteration, as

$$(32)\qquad \beta^{i+1} = \beta^i + (\tilde{x}'\tilde{x})^{-1}\tilde{x}'\tilde{e}.$$

Thus, an ordinary least-squares program can again perform the scoring iteration, and $(\tilde{x}'\tilde{x})^{-1}$ from this calculation will be the final variance-covariance matrix of the maximum likelihood estimates of β.

Under the conditions of Crowder's [7] theorem for martingales, it can be established that the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$ are asymptotically normally distributed with limiting distribution

$$(33)\qquad \begin{aligned} &\sqrt{T}(\hat{\alpha} - \alpha) \xrightarrow{D} N\left(0, \mathcal{I}_{\alpha\alpha}^{-1}\right), \\ &\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{D} N\left(0, \mathcal{I}_{\beta\beta}^{-1}\right). \end{aligned}$$

7. GAINS IN EFFICIENCY FROM MAXIMUM LIKELIHOOD ESTIMATION

The gain in efficiency from using the maximum-likelihood estimation rather than OLS has been asserted above. In this section, the gains are calculated for a special case. Consider the linear stationary ARCH model with $p = 1$ and all x_t exogenous. This is the case where the Gauss–Markov theorem applies and OLS has a variance matrix $\sigma^2(x'x)^{-1} = E\epsilon_t^2(\sum_t x_t'x_t)^{-1}$. The stationary variance is $\sigma^2 = \alpha_0/(1 - \alpha_1)$.

The information matrix for this case becomes, from (25),

$$E\left[\sum_t x_t'x_t\left(h_t^{-1} + 2\epsilon_t^2\alpha_1^2/h_{t+1}^2\right)\right].$$

With x exogenous, the expectation is only necessary over the scale factor. Because the disturbance process is stationary, the variance-covariance matrix is proportional to that for OLS and the relative efficiency depends only upon the scale factors. The relative efficiency of MLE to OLS is, therefore,

$$R = E(h_t^{-1} + 2\epsilon_t^2\alpha_1^2/h_{t+1}^2)\sigma^2.$$

Now substitute $h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2$, $\sigma^2 = \alpha_0/1 - \alpha_1$, and $\gamma = \alpha_1/1 - \alpha_1$. Recognizing that ϵ_{t-1}^2 and ϵ_t^2 have the same density, define for each

$$u = \epsilon\sqrt{(1 - \alpha_1)/\alpha_0}\,.$$

The expression for the relative efficiency becomes

$$R = E\left(\frac{1+\gamma}{1+\gamma u^2}\right) + 2\gamma^2 E\frac{u^2}{(1+\gamma u^2)^2}, \tag{34}$$

where u has variance one and mean zero. From Jensen's inequality, the expected value of a reciprocal exceeds the reciprocal of the expected value and, therefore, the first term is greater than unity. The second is positive, so there is a gain in efficiency whenever $\gamma \neq 0$. Eu^{-2} is infinite because u^2 is conditionally chi squared with one degree of freedom. Thus, the limit of the relative efficiency goes to infinity with γ:

$$\lim_{\gamma\to\infty} R \to \infty.$$

For α_1 close to unity, the gain in efficiency from using a maximum likelihood estimator may be very large.

8. TESTING FOR ARCH DISTURBANCES

In the linear regression model, with or without lagged-dependent variables, OLS is the appropriate procedure if the disturbances are not conditionally heteroscedastic. Because the ARCH model requires iterative procedures, it may be desirable to test whether it is appropriate before going to the effort to estimate it. The Lagrange multiplier test procedure is ideal for this as in many similar cases. See, for example, Breusch and Pagan [**4**, **5**], Godfrey [**12**], and Engle [**9**].

Under the null hypothesis, $\alpha_1 = \alpha_2 \cdots = \alpha_p = 0$. The test is based upon the score under the null and the information matrix under the null. Consider the ARCH model with $h_t = h(z_t\alpha)$, where h is some differentiable function which, therefore, includes both the linear and exponential cases as well as lots of others and $z_t = (1, e_{t-1}^2, \ldots, e_{t-p}^2)$ where e_t are the ordinary least squares residuals. Under the null, h_t is a constant denoted h^0. Writing $\partial h_t/\partial\alpha = h'z_t'$, where h' is the scalar derivative of h, the score and information can be written as

$$\frac{\partial l}{\partial \alpha}\Big|_0 = \frac{h'}{2h^0}\sum_t z_t'\left(\frac{e_t^2}{h^0} - 1\right) = \frac{h^{0\prime}}{2h^0} z' f^0,$$

$$\mathcal{I}_{\alpha\alpha}^0 = \frac{1}{2}\left(\frac{h^{0\prime}}{h^0}\right)^2 E z' z,$$

and, therefore, the LM test statistic can be consistently estimated by

$$\xi^* = \tfrac{1}{2} f^{0\prime} z(z'z)^{-1} z' f^0 \tag{35}$$

where $z' = (z_1', \ldots, z_T')$, f^0 is the column vector of

$$\left(\frac{e_t^2}{h^0} - 1\right).$$

This is the form used by Breusch and Pagan [**4**] and Godfrey [**12**] for testing for heteroscedasticity. As they point out, all reference to the h function has disappeared and, thus, the test is the same for any h which is a function only of $z_t\alpha$.

In this problem, the expectation required in the information matrix could be evaluated quite simply under the null; this could have superior finite sample performance. A second simplification, which is appropriate for this model as well as the heteroscedasticity model, is to note that plim $f^{0\prime}f^0/T = 2$ because normality has already been assumed. Thus, an asymptotically equivalent statistic would be

$$\xi = T f^{0\prime} z(z'z)^{-1} z' f^0 / f^{0\prime} f^0 = TR^2 \tag{36}$$

where R^2 is the squared multiple correlation between f^0 and z. Since adding a constant and multiplying by a scalar will not change the R^2 of a regression, this is also the R^2 of the regression of e_t^2 on an intercept and p lagged values of e_t^2. The statistic will be asymptotically distributed as chi square with p degrees of freedom when the null hypothesis is true.

The test procedure is to run the OLS regression and save the residuals. Regress the squared residuals on a constant and p lags and test TR^2 as a χ_p^2. This will be an asymptotically locally most powerful test, a characterization it shares with likelihood ratio and Wald tests. The same test has been proposed by Granger and Anderson [**13**] to test for higher moments in bilinear time series.

9. ESTIMATION OF THE VARIANCE OF INFLATION

Economic theory frequently suggests that economic agents respond not only to the mean, but also to higher moments of economic random variables. In financial theory, the variance as well as the mean of the rate of return are determinants of portfolio decisions. In macroeconomics, Lucas [**16**], for example, argues that the variance of inflation is a determinant of the response to various shocks. Furthermore, the variance of inflation may be of independent interest as it is the unanticipated component which is responsible for the bulk of the welfare loss due to inflation. Friedman [11] also argues that, as high inflation will generally be associated with high variability of inflation, the statistical relationship between inflation and unemployment should have a positive slope, not a negative one as in the traditional Phillips curve.

Measuring the variance of inflation over time has presented problems to various researchers. Khan [14] has used the absolute value of the first difference of inflation while Klein [15] has used a moving variance around a moving mean. Each of these approaches makes very simple assumptions about the mean of the distribution, which are inconsistent with conventional econometric approaches. The ARCH method allows a conventional regression specification for the mean function, with a variance which is permitted to change stochastically over the sample period. For a comparison of several measures for U.S. data, see Engle [10].

A conventional price equation was estimated using British data from 1958-II through 1977-II. It was assumed that price inflation followed wage increases; thus the model is a restricted transfer function.

Letting $\dot{p}$ be the first difference of the log of the quarterly consumer price index and w be the log of the quarterly index of manual wage rates, the model chosen after some experimentation was

$$\dot{p} = \beta_1 \dot{p}_{-1} + \beta_2 \dot{p}_{-4} + \beta_3 \dot{p}_{-5} + \beta_4 (p - w)_{-1} + \beta_5. \tag{37}$$

The model has typical seasonal behavior with the first, fourth, and fifth lags of the first difference. The lagged value of the real wage is the error correction mechanism of Davidson, et al. [8], which restricts the lag weights to give a constant real wage in the long run. As this is a reduced form, the current wage rate cannot enter.

The least squares estimates of this model are given in Table I. The fit is quite good, with less than 1 per cent standard error of forecast, and all t statistics greater than 3. Notice that $\dot{p}_{-4}$ and $\dot{p}_{-5}$ have equal and opposite signs, suggesting that it is the acceleration of inflation one year ago which explains much of the short-run behavior in prices.

TABLE I

ORDINARY LEAST SQUARES (36)[a]

Variable	$\dot{p}_{-1}$	$\dot{p}_{-4}$	$\dot{p}_{-5}$	$(p-w)_{-1}$	Const	$\alpha_0\ (\times 10^{-6})$	α_1
Coeff.	0.334	0.408	−0.404	−0.0559	0.0257	89	0
St Err.	0.103	0.110	0.114	0.0136	0.00572		
t Stat.	3.25	3.72	3.55	4.12	4.49		

[a] Dependent variable $p = \log(P) - \log(P_{-1})$ where P is quarterly U.K. consumer price index $w = \log(W)$ where W is the U.K. index of manual wage rates. Sample period 1958-II to 1977-II.

To establish the reliability of the model by conventional criteria, it was tested for serial correlation and for coefficient restrictions. Godfrey's [12] Lagrange multiplier test, for serial correlation up to sixth order, yields a chi-squared statistic with 6 degrees of freedom of 4.53, which is not significant, and the square of Durbin's h is 0.57. Only the 9th autocorrelation of the least squares residuals exceeds two asymptotic standard errors and, thus, the hypothesis of white noise disturbances can be accepted. The model was compared with an unrestricted regression, including all lagged p and w from one quarter through six. The asymptotic F statistic was 2.04, which is not significant at the 5 per cent level. When (37) was tested for the exclusion of w_{-1} through w_{-6}, the statistic was 2.34, which is barely significant at the 5 per cent but not the 2.5 per cent level. The only variable which enters significantly in either of these regressions is w_{-6} and it seems unattractive to include this alone.

The Lagrange multiplier test for a first-order linear ARCH effect for the model in (37) was not significant. However, testing for a fourth-order linear ARCH process, the chi-squared statistic with 4 degrees of freedom was 15.2, which is highly significant. Assuming that agents discount past residuals, a linearly declining set of weights was formulated to give the model

$$h_t = \alpha_0 + \alpha_1(0.4\epsilon_{t-1}^2 + 0.3\epsilon_{t-2}^2 + 0.2\epsilon_{t-3}^2 + 0.1\epsilon_{t-4}^2) \tag{38}$$

which is used in the balance of the paper. A two-parameter variance function was chosen because it was suspected that the nonnegativity and stationarity constraints on the α's would be hard to satisfy in an unrestricted model. The chi-squared test for $\alpha_1 = 0$ in (38) was 6.1, which has one degree of freedom.

One step of the scoring algorithm was employed to estimate model (37) and (38). The scoring step on α was performed first and then, using the new efficient $\hat{\alpha}$, the algorithm obtains in one step, efficient estimates of β. These are given in Table II. The procedure was also iterated to convergence by doing three steps on α, followed by three steps on β, followed by three more steps on α, and so forth. Convergence, within 0.1 per cent of the final value, occurred after two sets of α and β steps. These results are given in Table III.

The maximum likelihood estimates differ from the least squares effects primarily in decreasing the sizes of the short-run dynamic coefficients and increasing

TABLE II

MAXIMUM LIKELIHOOD ESTIMATES OF ARCH MODEL (36) (37)
ONE-STEP SCORING ESTIMATES[a]

Variable	$\dot{p}_{-1}$	$\dot{p}_{-4}$	$\dot{p}_{-5}$	$(p-w)_{-1}$	Const.	$\alpha_0\ (\times 10^{-6})$	α_1
Coeff.	0.210	0.270	−0.334	−0.0697	0.0321	19	0.846
St. Err.	0.110	0.094	0.109	0.0117	0.00498	14	0.243
t Stat.	1.90	2.86	3.06	5.98	6.44	1.32	3.49

[a] Dependent variable $p = \log(P) - \log(P_{-1})$ where P is quarterly U.K. consumer price index, $w = \log(W)$ where W is the U.K. index of manual wage rates. Sample period 1958-II to 1977-II.

TABLE III
MAXIMUM LIKELIHOOD ESTIMATES OF ARCH MODEL (36) (37)
ITERATED ESTIMATES[a]

Variables	p_{-1}	p_{-4}	$\dot{p}_{-5}$	$(p-w)_{-1}$	Const	$\alpha_0\ (\times 10^{-6})$	α_1
Coeff.	0.162	0.264	−0.325	−0.0707	0.0328	14	0.955
St. Err.	0.108	0.0892	0.0987	0.0115	0.00491	8.5	0.298
t Stat.	1.50	2.96	3.29	6.17	6.67	1.56	3.20

[a] Dependent variable $p = \log(P) - \log(P_{-1})$ where P is quarterly U K consumer price index $w = \log(W)$ where W is the U K index of manual wage rates Sample period 1958-II to 1977-II

the coefficient on the long run, as incorporated in the error correction mechanism. The acceleration term is not so clearly implied as in the least squares estimates. These seem reasonable results, since much of the inflationary dynamics are estimated by a period of very severe inflation in the middle seventies. This, however, is also the period of the largest forecast errors and, hence, the maximum likelihood estimator will discount these observations. By the end of the sample period, inflationary levels were rather modest and one might expect that the maximum likelihood estimates would provide a better forecasting equation.

The standard errors for ordinary least squares are generally greater than for maximum likelihood. The least squares standard errors are 15 per cent to 25 per cent greater, with one exception where the standard error actually falls by 5 per cent to 7 per cent. As mentioned earlier, however, the least squares estimates are biased when there are lagged dependent variables. The Wald test for $\alpha_1 = 0$ is also significant.

The final estimates of h_t are the one-step-ahead forecast variances. For the one-step scoring estimator, these vary from 23×10^{-6} to 481×10^{-6}. That is, the forecast standard deviation ranges from 0.5 per cent to 2.2 per cent, which is more than a factor of 4. The average of the h_t, since 1974, is 230×10^{-6}, as compared with 42×10^{-6} during the last four years of the sixties. Thus, the standard deviation of inflation increased from 0.6 per cent to 1.5 per cent over a few years, as the economy moved from the rather predictable sixties into the chaotic seventies.

In order to determine whether the confidence intervals arising from the ARCH model were superior to the least squares model, the outliers were examined. The expected number of residuals exceeding two (conditional) standard deviations is 3.5. For ordinary least squares, there were 5 while ARCH produced 3. For least squares these occurred in '74-I, '75-I, '75-II, '75-IV, and '76-II; they all occur within three years of each other and, in fact, three of them are in the same year. For the ARCH model, they are much more spread out and only one of the least squares points remains an outlier, although the others are still large. Examining the observations exceeding one standard deviation shows similar effects. In the seventies, there were 13 OLS and 12 ARCH residuals outside one sigma, which are both above the expected value of 9. In the sixties, there were 6 for OLS, 10 for ARCH and an expected number of 12. Thus, the number of outliers for ordinary least squares is reasonable; however, the timing of their occurrence is far from random. The ARCH model comes closer to truly random residuals after standardizing for their conditional distributions.

This example illustrates the usefulness of the ARCH model for improving the performance of a least squares model and for obtaining more realistic forecast variances.

University of California, San Diego

Manuscript received July, 1979; final revision received July, 1981.

APPENDIX

PROOF OF THEOREM 1: Let

(A2) $$w_t' = (y_t^{2r}, y_t^{2(r-1)}, \dots, y_t^2).$$

First, it is shown that there is an upper triangular $r \times r$ matrix A and $r \times 1$ vector b such that

(A2) $$E(w_t \mid \psi_{t-1}) = b + Aw_{t-1}.$$

For any zero-mean normal random variable u, with variance σ^2,

$$E(u^{2r}) = \sigma^{2r} \prod_{j=1}^{r} (2j-1)$$

Because the conditional distribution of y is normal

(A3) $$E(y_t^{2m} \mid \psi_{t-1}) = h_t^{2m} \prod_{j=1}^{m} (2j-1)$$

$$= (\alpha_1 y_{t-1}^2 + \alpha_0)^m \prod_{j=1}^{m} (2j-1).$$

Expanding this expression establishes that the moment is a linear combination of w_{t-1} Furthermore, only powers of y less than or equal to $2m$ are required, therefore, A in (A2) is upper triangular

Now

$$E(w_t \mid \psi_{t-2}) = b + A(b + Aw_{t-2})$$

or in general

$$E(w_t \mid \psi_{t-k}) = (I + A + A^2 + \cdots + A^{k-1})b + A^k w_{t-k}.$$

Because the series starts indefinitely far in the past with $2r$ finite moments, the limit as k goes to infinity exists if, and only if, all the eigenvalues of A lie within the unit circle.

The limit can be written as

$$\lim_{k \to \infty} E(w_t \mid \psi_{t-k}) = (I - A)^{-1} b,$$

which does not depend upon the conditioning variables and does not depend upon t. Hence, this is an expression for the stationary moments of the unconditional distribution of y.

(A4) $$E(w_t) = (I - A)^{-1} b.$$

It remains only to establish that the condition in the theorem is necessary and sufficient to have all eigenvalues lie within the unit circle. As the matrix has already been shown to be upper triangular, the diagonal elements are the eigenvalues. From (A3), it is seen that the diagonal elements are simply

$$\alpha_1^m \prod_{j=1}^{m} (2j-1) = \prod_{j=1}^{m} \alpha_1(2j-1) \equiv \theta_m$$

for $m = 1, \ldots, r$. If θ_r exceeds or equals unity, the eigenvalues do not lie in the unit circle. It must also be shown that if $\theta_r < 1$, then $\theta_m < 1$ for all $m < r$. Notice that θ_m is a product of m factors which are monotonically increasing. If the mth factor is greater than one, then θ_{m-1} will necessarily be smaller than θ_m. If the mth factor is less than one, all the other factors must also be less than one and, therefore, θ_{m-1} must also have all factors less than one and have a value less than one. This establishes that a necessary and sufficient condition for all diagonal elements to be less than one is that $\theta_r < 1$, which is the statement in the theorem. *Q.E.D.*

Proof of Theorem 2: Let

$$w_t' = \left(y_t^2, y_{t-1}^2, \ldots, y_{t-p}^2\right).$$

Then in terms of the companion matrix A,

$$E(w_t \mid \psi_{t-1}) = b + Aw_{t-1} \tag{A5}$$

where $b' = (\alpha_0, 0, \ldots, 0)$ and

$$A = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_p & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

Taking successive expectations

$$E(w_t \mid \psi_{t-k}) = (I + A + A^2 + \cdots + A^{k-1})b + A^k w_{t-k}.$$

Because the series starts indefinitely far in the past with finite variance, if, and only if, all eigenvalues lie within the unit circle, the limit exists and is given by

$$\lim_{k\to\infty} E(w_t \mid \psi_{t-k}) = (I - A)^{-1} b \tag{A6}$$

As this does not depend upon initial conditions or on t, this vector is the common variance for all t. As is well known in time series analysis, this condition is equivalent to the condition that all the roots of the characteristic equation, formed from the α's, lie outside the unit circle. See Anderson [2, p. 177]. Finally, the limit of the first element can be rewritten as

$$Ey_t^2 = \alpha_0 \Big/ \left(1 - \sum_{j=1}^{p} \alpha_j\right). \tag{A7}$$

Q.E.D.

Proof of Theorem 3: Clearly, under the conditions, $h(\xi_t) \geq \alpha_0 > 0$, establishing part (a). Let

$$\begin{aligned}\phi_{i,m,t} &= E\left(|\partial h(\xi_t)/\partial \alpha_i|\,|\partial h(\xi_t)/\partial \xi_{t-m}| \mid \psi_{t-m-1}\right)\\ &= 2\alpha_m E\left(|\xi_{t-i}|^2 |\xi_{t-m}| \psi_{t-m-1}\right).\end{aligned}$$

Now there are three cases; $i > m$, $i = m$, and $i < m$. If $i > m$, then $\xi_{t-i} \in \psi_{t-m-1}$ and the conditional expectation of $|\xi_{t-m}|$ is finite, because the conditional density is normal. If $i = m$, then the expectation becomes $E(|\xi_{t-m}|^3 \mid \psi_{t-m-1})$. Again, because the conditional density is normal, all moments exist including the expectation of the third power of the absolute value. If $i < m$, the expectation is taken in two parts, first with respect to $t - i - 1$:

$$\phi_{i,m,t} = 2\alpha_m E\left\{ |\xi_{t-m}| E(\xi_{t-i}^2 \mid \psi_{t-i-1}) \mid \psi_{t-m-1} \right\}$$

$$= 2\alpha_m E\left\{ |\xi_{t-m}| \alpha_0 + \sum_{j=1}^{p} \alpha_j \xi_{t-i-j}^2) \mid \psi_{t-m-1} \right\}$$

$$= 2\alpha_m \alpha_0 E\{ \xi_{t-m} \mid \psi_{t-m-1} \} + \sum_{j=1}^{p} \alpha_j \phi_{i+j,m,t}.$$

In the final expression, the initial index on ϕ is larger and, therefore, may fall into either of the preceding cases, which, therefore, establishes the existence of the term. If there remain terms with $i + j < m$, the recursion can be repeated. As all lags are finite, an expression for $\phi_{i,m,t}$ can be written as a constant times the third absolute moment of ξ_{t-m} conditional on ψ_{t-m-1}, plus another constant times the first absolute moment. As these are both conditionally normal, and as the constants must be finite as they have a finite number of terms, the second part of the regularity condition has been established *Q.E.D.*

To establish Theorem 4, a careful symmetry argument is required, beginning with the following lemma.

LEMMA: *Let u and v be any two random variables. $E(g(u,v) \mid v)$ will be an anti-symmetric function of v if g is anti-symmetric in v, the conditional density of $u \mid v$ is symmetric in v, and the expectation exists.*

PROOF:

$$E(g(u,-v) \mid -v) = -E(g(u,v) \mid -v) \qquad \text{because } g \text{ is anti-symmetric in } v$$

$$= -E(g(u,v) \mid v) \qquad \text{because the conditional density is symmetric.}$$

Q.E.D.

PROOF OF THEOREM 4: The i, j element of $I_{\alpha\beta}$ is given by

$$(I_{\alpha\beta})_{ij} = \frac{1}{2T} \sum_t E\left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \beta_j} \right)$$

$$= -\frac{1}{2T} \sum_t \sum_{m=1}^{p} E\left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \epsilon_{t-m}} x_{jt-m} \right] \qquad \text{by the chain rule.}$$

If the expectation of the term in square brackets, conditional on ψ_{t-m-1}, is zero for all i, j, t, m, then the theorem is proven

$$E\left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \epsilon_{t-m}} x_{jt-m} \mid \psi_{t-m-1} \right) = x_{jt-m} E\left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \epsilon_{t-m}} \mid \psi_{t-m-1} \right)$$

because x_{jt-m} is either exogenous or it is a lagged dependent variable, in which case it is included in ψ_{t-m-1}.

$$\left| E\left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \epsilon_{t-m}} \mid \psi_{t-m-1} \right) \right| \le E\left(\frac{1}{h_t^2} \left| \frac{\partial h_t}{\partial \alpha_i} \right| \left| \frac{\partial h_t}{\partial \epsilon_{t-m}} \right| \mid \psi_{t-m-1} \right)$$

$$\le \frac{1}{\delta^2} E\left(\left| \frac{\partial h_t}{\partial \alpha_i} \right| \left| \frac{\partial h_t}{\partial \epsilon_{t-m}} \right| \mid \psi_{t-m-1} \right)$$

by part (a) of the regularity conditions and this integral is finite by part (b) of the condition. Hence, each term is finite. Now take the expectation in two steps, first with respect to ψ_{t-m}. This must therefore also be finite.

$$E\left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \epsilon_{t-m}} \mid \psi_{t-m} \right) \equiv g(\epsilon_{t-m}).$$

By the symmetry assumption, h_t^{-1} is symmetric in ϵ_{t-m}, $\partial h_t / \partial \epsilon_{t-m}$ is anti-symmetric. Therefore, the whole expression is anti-symmetric in ϵ_{t-m}, which is part of the conditioning set ψ_{t-m}. Because h is symmetric, the conditional density must be symmetric in ϵ_{t-m} and the lemma can be invoked to show that $g(\epsilon_{t-m})$ is anti-symmetric.

Finally, taking expectations of g conditional on ψ_{t-m-1} gives zero, because the density of ϵ_{t-m} conditional on the past is a symmetric (normal) density and the theorem is established. *Q.E.D.*

REFERENCES

[1] AMEMIYA, T.: "Regression Analysis when the Variance of the Dependent Variable is Proportional to the Square of Its Expectation," *Journal of the American Statistical Association*, 68(1973), 928–934.

[2] ANDERSON, T. W.: *The Statistical Analysis of Time Series*. New York: John Wiley and Sons, 1958.

[3] BELSLEY, DAVID: "On the Efficient Computation of Non-Linear Full-Information Maximum Likelihood Estimator," paper presented to the European Meetings of the Econometric Society, Athens, 1979.

[4] BREUSCH, T. S., AND A. R. PAGAN: "A Simple Test for Heteroscedasticity and Random Coefficient Variation," *Econometrica*, 46(1978), 1287–1294.

[5] ———: "The Lagrange Multiplier Test and Its Applications to Model Specification," *Review of Economic Studies*, 47(1980), 239–254.

[6] COX, D. R., AND D. V. HINKLEY: *Theoretical Statistics*. London: Chapman and Hall, 1974.

[7] CROWDER, M. J.: "Maximum Likelihood Estimation for Dependent Observations," *Journal of the Royal Statistical Society*, Series B, 38(1976), 45–53.

[8] DAVIDSON, J. E. H., D. F. HENDRY, F. SRBA, AND S. YEO: "Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom," *The Economic Journal*, 88(1978), 661–691.

[9] ENGLE, R. F.: "A General Approach to the Construction of Model Diagnostics Based upon the Lagrange Multiplier Principle," University of California, San Diego Discussion Paper 79-43, 1979.

[10] ———: "Estimates of the Variance of U.S. Inflation Based on the ARCH Model," University of California, San Diego Discussion Paper 80-14, 1980.

[11] FRIEDMAN, MILTON: "Nobel Lecture: Inflation and Unemployment," *Journal of Political Economy*, 85(1977), 451–472.

[12] GODFREY, L. G.: "Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables," *Econometrica*, 46(1978), 1293–1302.

[13] GRANGER, C. W. J., AND A. ANDERSEN: *An Introduction to Bilinear Time-Series Models*. Göttingen: Vandenhoeck and Ruprecht, 1978.

[14] KHAN, M. S.: "The Variability of Expectations in Hyperinflations," *Journal of Political Economy*, 85(1977), 817–827.

[15] KLEIN, B.: "The Demand for Quality-Adjusted Cash Balances: Price Uncertainty in the U.S. Demand for Money Function," *Journal of Political Economy*, 85(1977), 692–715.

[16] LUCAS, R. E., JR.: "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, 63(1973), 326–334.

[17] MCNEES, S. S.: "The Forecasting Record for the 1970's," *New England Economic Review*, September/October 1979, 33–53.

[18] WHITE, H.: "A Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica*, 48(1980), 817–838.

DUFF & PHELPS

2020 SBBI® Yearbook

Stocks, Bonds, Bills, and Inflation®

U.S. Capital Markets Performance by
Asset Class 1926–2019

Appendix A

Appendix A

Appendix A-1

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year	Jan - Dec
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Appendix A-7

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year	Jan-Dec

Appendix A-7

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year	Jan-Dec

Appendix A-7

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year	Jan–Dec
1988	[illegible]	[illegible]	[illegible]	0.0070	[illegible]	0.0076	0.0071	0.0083	[illegible]	0.0076	0.0070	0.007[illegible]	1988	[illegible]
1989	0.0036	[illegible]	[illegible]	0.0070	0.0060	0.0070	0.0068	0.006[illegible]	0.006[illegible]	0.0072	0.0054	0.0064	1989	[illegible]
1990	[illegible]	[illegible]	[illegible]	[illegible]	0.0075	0.0068	0.00[illegible]4	0.0071	0.0069	0.0081	0.0071	0.00[illegible]	1990	[illegible]
1991	[illegible]	[illegible]	[illegible]	[illegible]	0.0068	[illegible]	0.0076	0.0063	0.0068	0.0065	0.0056	0.006[illegible]	1991	[illegible]
1992	[illegible]	[illegible]	[illegible]	[illegible]	0.0061	0.0067	0.0063	0.0060	0.0058	0.005[illegible]	0.0061	0.006[illegible]	1992	[illegible]
1993	[illegible]	[illegible]	[illegible]	0.0057	[illegible]	0.0062	0.0054	0.0056	0.0050	0.0043	0.0053	0.005[illegible]	1993	[illegible]
1994	[illegible]	[illegible]	[illegible]	[illegible]	0.0063	0.0061	0.0060	0.0066	0.005[illegible]	0.0056	0.0054	0.00[illegible]	1994	[illegible]
1995	0.005[illegible]	[illegible]	[illegible]	[illegible]	0.0066	0.0054	0.0056	0.005[illegible]	0.0052	0.0057	0.0051	0.004[illegible]	1995	[illegible]
1996	[illegible]	[illegible]	[illegible]	[illegible]	0.0052	0.0054	0.0052	0.0057	0.0060	0.0058	0.005[illegible]	0.005[illegible]	1996	[illegible]
1997	0.0054	0.0048	[illegible]	[illegible]	0.0053	0.0059	0.0058	0.0049	0.0050	0.005[illegible]	0.0047	0.005[illegible]	1997	[illegible]
1998	0.005[illegible]	[illegible]	[illegible]	0.0049	0.0049	0.0052	0.0049	0.0048	0.0044	0.004[illegible]	0.004[illegible]	0.004[illegible]	1998	[illegible]
1999	0.0049	0.0044	[illegible]	0.0046	0.0045	0.0055	0.0051	0.0054	0.005[illegible]	0.0050	0.0056	0.005[illegible]	1999	[illegible]
2000	0.0042	0.0040	[illegible]	[illegible]	0.0036	0.0052	0.0052	0.0050	0.0046	0.0053	0.0048	0.0045	2000	[illegible]
2001	[illegible]	[illegible]	[illegible]	0.0041	0.0050	0.004[illegible]	0.005[illegible]	0.0046	0.0043	0.0043	[illegible]	0.0046	2001	[illegible]
2002	0.0016	[illegible]	[illegible]	[illegible]	0.0040	0.0041	0.005[illegible]	0.0044	0.004[illegible]	0.0046	0.00[illegible]	0.004[illegible]	2002	[illegible]
2003	0.0042	[illegible]	[illegible]	[illegible]	[illegible]	0.0030	0.0030	0.0042	0.0036	0.0041	0.00[illegible]	0.003[illegible]	2003	[illegible]
2004	[illegible]	[illegible]	[illegible]	[illegible]	0.0046	0.0041	0.0045	0.0045	0.00[illegible]	0.0038	0.0041	0.004[illegible]	2004	[illegible]
2005	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0036	0.0034	0.0040	0.00[illegible]	0.00[illegible]9	0.0039	0.0038	2005	[illegible]
2006	[illegible]	[illegible]	[illegible]	[illegible]	0.0049	0.0044	0.004[illegible]	0.0043	0.00[illegible]	0.0042	0.00[illegible]	0.005[illegible]	2006	[illegible]
2007	[illegible]	[illegible]	[illegible]	0.0042	0.0041	0.0040	0.0046	[illegible]	0.003[illegible]	0.004[illegible]	0.00[illegible]9	0.00[illegible]	2007	[illegible]
2008	0.0043	[illegible]	[illegible]	[illegible]	0.0025	0.0043	0.0039	0.0056	0.00[illegible]4	0.003[illegible]	0.00[illegible]	0.00[illegible]	2008	[illegible]
2009	0.0040	[illegible]	[illegible]	[illegible]	[illegible]	0.0038	0.0035	0.0036	0.0034	0.00[illegible]	0.00[illegible]	0.00[illegible]4	2009	[illegible]
2010	[illegible]	[illegible]	[illegible]	[illegible]	0.0024	0.0031	0.0031	0.0032	0.002[illegible]	0.00[illegible]	0.0022	0.00[illegible]	2010	[illegible]
2011	[illegible]	[illegible]	[illegible]	[illegible]	0.0036	0.002[illegible]	0.002[illegible]	0.0034	0.002[illegible]	0.002[illegible]	0.00[illegible]	0.00[illegible]	2011	[illegible]
2012	[illegible]	[illegible]	[illegible]	0.0026	0.0023	0.0018	0.0020	0.0018	0.001[illegible]	0.0021	0.00[illegible]	0.0019	2012	[illegible]
2013	0.002[illegible]	[illegible]	[illegible]	0.0026	0.0023	0.00[illegible]4	0.0030	0.0023	0.0023	0.0029	0.00[illegible]	0.0031	2013	[illegible]
2014	[illegible]	[illegible]	[illegible]	0.0023	0.0023	0.0025	0.002[illegible]	0.0026	0.002[illegible]	0.0025	0.002[illegible]	0.00[illegible]	2014	[illegible]
2015	[illegible]	0.0026	[illegible]	[illegible]	0.0020	0.0022	0.0024	0.0022	0.00[illegible]	0.0021	0.00[illegible]	0.00[illegible]	2015	[illegible]
2016	[illegible]	[illegible]	[illegible]	[illegible]	0.0020	0.0018	0.0014	0.0016	0.0015	0.001[illegible]	0.0018	0.00[illegible]	2016	[illegible]
2017	[illegible]	[illegible]	0.0018	[illegible]	0.0021	0.0021	0.00[illegible]	0.0022	0.001[illegible]	0.00[illegible]	0.00[illegible]	0.0026	2017	[illegible]
2018	0.0024	[illegible]	[illegible]	0.0021	0.0018	0.0022	0.0025	0.0025	0.002[illegible]	0.0030	0.002[illegible]	0.00[illegible]	2018	[illegible]
2019	0.0023	[illegible]	[illegible]	[illegible]	0.0022	0.001[illegible]	[illegible]	[illegible]	0.001[illegible]	0.0016	[illegible]	[illegible]	2019	[illegible]

[illegible]

U.S. Capital Markets Performance
by Asset Class 1926–2019

2020
SBBI® Yearbook

Stocks, Bonds, Bills, and Inflation

Roger G. Ibbotson

DUFF & PHELPS

20-Year vs. 30-Year Treasuries

The U.S. Treasury periodically changes the maturities that it issues. For example, in April 1986 the U.S. Treasury stopped issuing 20-year Treasuries, and from October 2001 through January 2006 the U.S. Treasury did not issue 30-year bonds (it resumed issuing 30-year Treasury bonds in February 2006), making the 10-year bond the longest term Treasury security issued over the October 2001 – January 2006 period. Most recently, on January 16, 2020 the U.S. Department of the Treasury announced its plans to issue a 20-year nominal coupon bond in the first half of calendar year 2020, the first time a 20-year maturity will be offered since March 1986.[10.7,10.11]

Our methodology for estimating the long-horizon equity risk premium makes use of the income return on a 20-year Treasury bond. While a 30-year bond is theoretically more correct when dealing with the long-term nature of business valuation,[10.12] 30-year Treasury securities have an issuance history that is on-again-off-again. Ibbotson Associates creates a series of returns using bonds on the market with approximately 20 years to maturity because Treasury bonds of this maturity are available over a long history, while Treasury bonds of 30-years are not.

Income Return

Another point to keep in mind when calculating the equity risk premium is that the income return on the appropriate-horizon Treasury security, rather than the total return, is used in the calculation.

The total return comprises three return components: the income return, the capital appreciation return, and the reinvestment return. The income return is defined as the portion of the total return that results from a periodic cash flow or, in this case, the bond coupon payment. The capital appreciation return results from the price change of a bond over a specific period. Bond prices generally change in reaction to unexpected fluctuations in yields. Reinvestment return is the return on a given month's investment income when reinvested into the same asset class in the subsequent months of the year. The income return is thus used in the estimation of the equity risk premium because it represents the truly riskless portion of the return.

Arithmetic vs. Geometric Mean

The equity risk premium data presented in this book are arithmetic average risk premiums as opposed to geometric average risk premiums. The arithmetic average equity risk premium can be demonstrated to be most appropriate when discounting future cash flows. For use as the expected equity risk premium in either the CAPM or the building block approach, the arithmetic mean or the simple difference of the arithmetic means of stock market returns and riskless rates is the relevant number.

[illegible] Department of the Treasury [illegible]

[10.11] [illegible] "Treasury to Issue New 20-Year Bond in First Half of [illegible]" The Wall Street Journal, [illegible] 2020 [illegible]

[10.12] [illegible] cost of capital [illegible] expected return [illegible] required return [illegible] rate [illegible] **Note:** [illegible] platform [illegible] Cost of Capital Navigator [illegible] through the process of estimating the cost of capital [illegible] component [illegible] The Cost of Capital Navigator can be used to estimate [illegible] of equity capital [illegible] countries [illegible] For more information, visit [illegible]

10890

The Market Risk Premium: Expectational Estimates Using Analysts' Forecasts

Robert S. Harris and Felicia C. Marston

Using expectational data from financial analysts, we estimate a market risk premium for US stocks. Using the S&P 500 as a proxy for the market portfolio, the average market risk premium is found to be 7.14% above yields on long-term US government bonds over the period 1982-1998. This risk premium varies over time; much of this variation can be explained by either the level of interest rates or readily available forward-looking proxies for risk. The market risk premium appears to move inversely with government interest rates suggesting that required returns on stocks are more stable than interest rates themselves. [JEL: G31, G12]

The notion of a market risk premium (the spread between investor required returns on safe and average risk assets) has long played a central role in finance. It is a key factor in asset allocation decisions to determine the portfolio mix of debt and equity instruments. Moreover, the market risk premium plays a critical role in the Capital Asset Pricing Model (CAPM), the most widely used means of estimating equity hurdle rates by practitioners. In recent years, the practical significance of estimating such a market premium has increased as firms, financial analysts, and investors employ financial frameworks to analyze corporate and investment performance. For instance, the increased use of Economic Value Added (EVA) to assess corporate performance has provided a new impetus for estimating capital costs.

The most prevalent approach to estimating the market risk premium relies on some average of the historical spread between returns on stocks and bonds.[1] This choice has some appealing characteristics but is subject to many arbitrary assumptions such as the relevant period for taking an average. Compounding the difficulty of using historical returns is the well noted fact that standard models of consumer choice would predict much lower spreads between equity and debt returns than have occurred in US markets — the so called equity risk premium puzzle (see Welch, 2000 and Siegel and Thaler, 1997). In addition, theory calls for a forward-looking risk premium that could well change over time.

This paper takes an alternate approach by using expectational data to estimate the market risk premium. The approach has two major advantages for practitioners. First, it provides an independent estimate that can be compared to historical averages. At a minimum, this can help in understanding likely ranges for risk premia. Second, expectational data allow investigation of changes in risk premia over time. Such time variations in risk premia serve as important signals from investors that should affect a host of financial decisions. This paper provides new tests of whether changes in risk premia over time are linked to forward-looking measures of risk. Specifically, we look at the

Robert S. Harris is the C. Stewart Sheppard Professor of Business Administration and Felicia C. Marston is an Associate Professor at the University of Virginia, Charlottesville, VA 22906.

The authors thank Erik Benrud, an anonymous reviewer, and seminar participants at the University of Virginia, the University of Connecticut and at the SEC for comments. Thanks to Darden Sponsors, IVA, the Walker Family Fund, and McIntire Associates for support of this research and to IBES, Inc. for supplying data.

[1] Bruner, Eades, Harris, and Higgins (1998) provide survey evidence on both textbook advice and practitioner methods for estimating capital costs. As testament to the market for cost of capital estimates, Ibbotson Associates (1998) publishes a "Cost of Capital Quarterly."

relationship between the risk premium and four *ex-ante* measures of risk: the spread between yields on corporate and government bonds, consumer sentiment about future economic conditions, the average level of dispersion across analysts as they forecast corporate earnings, and the implied volatility on the S&P500 Index derived from options data.

Section I provides background on the estimation of equity required returns and a brief discussion of current practice in estimating the market risk premium. In Section II, models and data are discussed. Following a comparison of the results to historical returns in Section III, we examine the time-series characteristics of the estimated market premium in Section IV. Finally, conclusions are offered in Section V.

I. Background

The notion of a "market" required rate of return is a convenient and widely used construct. Such a rate (k) is the minimum level of expected return necessary to compensate investors for bearing the average risk of equity investments and receiving dollars in the future rather than in the present. In general, k will depend on returns available on alternative investments (e.g., bonds). To isolate the effects of risk, it is useful to work in terms of a market risk premium (rp), defined as

$$rp = k - i, \tag{1}$$

where i = required return for a zero risk investment.

Lacking a superior alternative, investigators often use averages of historical realizations to estimate a market risk premium. Bruner, Eades, Harris, and Higgins (1998) provide recent survey results on best practices by corporations and financial advisors. While almost all respondents used some average of past data in estimating a market risk premium, a wide range of approaches emerged. "While most of our 27 sample companies appear to use a 60+ year historical period to estimate returns, one cited a window of less than ten years, two cited windows of about ten years, one began averaging with 1960, and another with 1952 data" (p. 22). Some used arithmetic averages, and some used geometric. This historical approach requires the assumptions that past realizations are a good surrogate for future expectations and, as typically applied, that the risk premium is constant over time. Carleton and Lakonishok (1985) demonstrate empirically some of the problems with such historical premia when they are disaggregated for different time periods or groups of firms. Siegel (1999) cites additional problems of using historical returns and argues that equity premium estimates from past data are likely too high. As Bruner et al. (1998) point out, few respondents cited use of expectational data to supplement or replace historical returns in estimating the market premium.

Survey evidence also shows substantial variation in empirical estimates. When respondents gave a precise estimate of the market premium, they cited figures from 4% to over 7% (Bruner et al., 1998). A quote from a survey respondent highlights the range in practice: "In 1993, we polled various investment banks and academic studies on the issue as to the appropriate rate and got anywhere between 2 and 8%, but most were between 6% and 7.4%." (Bruner et al., 1998). An informal sampling of current practice also reveals large differences in assumptions about an appropriate market premium. For instance, in a 1999 application of EVA analysis, Goldman Sachs Investment Research specifies a market risk premium of "3% from 1994-1997 and 3.5% from 1998-1999E for the S&P Industrials" (Goldman Sachs, 1999). At the same time, an April 1999 phone call to Stern Stewart revealed that their own application of EVA typically employed a market risk premium of 6%. In its application of the CAPM, Ibbotson Associates (1998) uses a market risk premium of 7.8%. Not surprisingly, academics do not agree on the risk premium either. Welch (2000) surveyed leading financial economists at major universities. For a 30-year horizon, he found a mean risk premium of 7.1% but a range from 1.5% to 15% with an interquartile range of 2.4% (based on 226 responses).

To provide additional insight on estimates of the market premium, we use publicly available expectational data. This expectational approach employs the dividend growth model (hereafter referred to as the discounted cash flow (DCF) model) in which a consensus measure of financial analysts' forecasts (FAF) of earnings is used as a proxy for investor expectations. Earlier work has used FAF in DCF models[1] but generally has covered a span of only a few years due to data availability.

II. Models and Data

The simplest and most commonly used version of the DCF model is employed to estimate shareholders' required rate of return, k, as shown in Equation (2):

[1] See Malkiel (1982), Brigham, Vinson, and Shome (1985), Harris (1986), and Harris and Marston (1992). The DCF approach with analysts' forecasts has been used frequently in regulatory settings. Ibbotson Associates (1998) use a variant of the DCF model with forward-looking growth rates; however, they do this as a separate technique and not as part of the CAPM. For their CAPM estimates, they use historical averages for the market risk premium.

$$k = \left(\frac{D_1}{P_0} \right) + g, \qquad (2)$$

where D_1 = dividend per share expected to be received at time one, P_0 = current price per share (time 0), and g = expected growth rate in dividends per share.' A primary difficulty in using the DCF model is obtaining an estimate of g, since it should reflect market expectations of future performance. This paper uses published FAF of long-run growth in earnings as a proxy for g. Equation (2) can be applied for an individual stock or any portfolio of companies. We focus primarily on its application to estimate a market premium as proxied by the S&P500.

FAF comes from IBES Inc. The mean value of individual analysts' forecasts of five-year growth rate in EPS is used as the estimate of g in the DCF model. The five-year horizon is the longest horizon over which such forecasts are available from IBES and often is the longest horizon used by analysts. IBES requests "normalized" five-year growth rates from analysts in order to remove short-term distortions that might stem from using an unusually high or low earnings year as a base. Growth rates are available on a monthly basis.

Dividend and other firm-specific information come from COMPUSTAT. D_1 is estimated as the current indicated annual dividend times $(1+g)$. Interest rates (both government and corporate) are from Federal Reserve Bulletins and *Moody's Bond Record*. Exhibit 1 describes key variables used in the study. Data are used for all stocks in the *Standard and Poor's 500* stock (S&P500) index followed by IBES. Since five-year growth rates are first available from IBES beginning in 1982, the analysis covers the period from January 1982-December 1998.

The approach used is generally the same approach as used in Harris and Marston (1992). For each month, a market required rate of return is calculated using each dividend-paying stock in the S&P500 index for which data are available. As additional screens for reliability of data, in a given month we eliminate a firm if there are fewer than three analysts' forecasts or if the standard deviation around the mean forecast exceeds 20%. Combined, these two screens eliminate fewer than 20 stocks a month. Later we report on the sensitivity of the results to various screens. The DCF model in Equation (2) is applied to each stock and the results weighted by market value of equity to produce the market-required return. The risk premium is constructed by subtracting the interest rate on government bonds.

We weighted 1998 results by year-end 1997 market values since the monthly data on market value did not extend through this period. Since data on firm-specific dividend yields were not available for the last four months of 1998 at the time of this study, the market dividend yield for these months was estimated using the dividend yield reported in the *Wall Street Journal* scaled by the average ratio of this figure to the dividend yield for our sample as calculated in the first eight months of 1998. Adjustments were then made using growth rates from IBES to calculate the market required return. We also estimated results using an average dividend yield for the month that employed the average of the price at the end of the current and prior months. These average dividend yield measures led to similar regression coefficients as those reported later in the paper.

For short-term horizons (quarterly and annual), past research (Brown, 1993) finds that on average analysts' forecasts are overly optimistic compared to realizations. However, recent research on quarterly horizons (Brown, 1997) suggests that analysts' forecasts for S&P500 firms do not have an optimistic bias for the period 1993-1996. There is very little research on the properties of five-year growth forecasts, as opposed to shorter horizon predictions. Boebel (1991) and Boebel, Harris, and Gultekin (1993) examine possible bias in analysts' five-year growth rates. These studies find evidence of optimism in IBES growth forecasts. In the most thorough study to date, Boebel (1991) reports that this bias seems to be getting smaller over time. His forecast data do not extend into the 1990s.

Analysts' optimism, if any, is not necessarily a problem for the analysis in this paper. If investors share analysts' views, our procedures will still yield unbiased estimates of required returns and risk premia. In light of the possible bias, however, we interpret the estimates as "upper bounds" for the market premium.

This study also uses four very different sources to create *ex ante* measures of equity risk at the market

Our methods follow Harris (1986) and Harris and Marston (1992) who discuss earlier research and the approach employed here, including comparisons of single versus multistage growth models. Since analysts' forecast growth in earnings per share, their projections should incorporate the anticipated effects of share repurchase programs. Dividends per share would grow at the same rate as EPS as long as companies manage a constant ratio of dividends to earnings on a per share basis. Based on S&P500 figures (see the Standard and Poor's website for their procedures), the ratio of DPS to EPS was .51 during the period 1982-89 and .52 for the period 1990-98. Lamdin (2001) discusses some issues if share repurchases destroy the equivalence of EPS and DPS growth rates. Theoretically, r is a risk-free rate, though its empirical proxy is only a "least risk" alternative that is itself subject to risk. For instance, Asness (2000) shows that over the 1946-1998 period bond volatility (in monthly realized returns) has increased relative to stock volatility, which would be consistent with a drop in the equity market premium.

Exhibit 1. Variable Definitions

k	=	Equity required rate return
P_t	=	Price per share
D_1	=	Expected dividend per share measured as current indicated annual dividend from COMPUSTAT multiplied by $(1 + g)$
g	=	Average financial analysts' forecast of five-year growth rate in earnings per share (from IBES)
r	=	Yield to maturity on long-term US government obligations (source Federal Reserve, 30-year constant maturity series)
rp	=	Equity risk premium calculated as $rp = k - r$
BSPREAD	=	Spread between yields on corporate and government bonds. BSPREAD = yield to maturity on long-term corporate bonds (Moody's average across bond rating categories) minus r
CON	=	Monthly consumer confidence index reported by the Conference Board (divided by 100)
DISP	=	Dispersion of analysts' forecasts at the market level
VOL	=	Volatility for the S+P500 index as implied by options data.

level. The first proxy comes from the bond market and is calculated as the spread between corporate and government bond yields (BSPREAD). The rationale is that increases in this spread signal investors' perceptions of increased riskiness of corporate activity that would be translated to both debt and equity owners. The second measure, CON, is the consumer confidence index reported by the Conference Board at the end of the month. While the reported index tends to be around 100, we rescale CON as the actual index divided by 100. We also examined use of CON as of the end of the prior month; however, in regression analysis, this lagged measure generally was not statistically significant in explaining the level of the market risk premium. The third measure, DISP, measures the dispersion of analysts' forecasts. Such analyst disagreement should be positively related to perceived risk since higher levels of uncertainty would likely generate a wider distribution of earnings forecasts for a given firm. DISP is calculated as the average of firm-specific standard deviations for each stock in the S&P500 covered by IBES. The firm-specific standard deviation is calculated based on the dispersion of individual analysts' growth forecasts around the mean of individual forecasts for that company in that month. DISP also was estimated using a value-weighted measure of analyst dispersion for the firms in our sample. The results reported use the equally weighted version but similar patterns were obtained with both constructions. Our final measure, VOL, is the implied volatility on the S&P500 index. As of the beginning of the month, a dividend-adjusted Black Scholes Formula is used to estimate the implied volatility in the S&P500 index option contract which expires on the third Friday of the month. The call premium, exercise price, and the level of the S&P500 index are taken from the *Wall Street Journal*, and treasury yields come from the Federal Reserve. Dividend yield comes from DRI. The option contract that is closest to being at the money is used.

III. Estimates of the Market Premium

Exhibit 2 reports both required returns and risk premia by year (averages of monthly data). The estimated risk premia are positive, consistent with equity owners demanding additional rewards over and above returns on debt securities. The average expectational risk premium (1982 to 1998) over

We examined two other proxies for Consumer Confidence. The Conference Board's Consumer Expectations Index yielded essentially the same results as those reported. The University of Michigan's Consumer Sentiment Indices tended to be less significantly linked to the market risk premium though coefficients were still negative.

For the regressions reported in Exhibit 6, the value-weighted dispersion measure actually exhibited more explanatory power. For regressions using the Prais-Winsten method (see footnote 7), the coefficient on DISP was not significant in 2 of the 4 cases.

Exhibit 2. Bond Market Yields, Equity Required Return, and Equity Risk Premium, 1982-1998

Values are averages of monthly figures in percent. i is the yield to maturity on long-term government bonds. k is the required return on the S&P500 estimated as a value weighted average using a discounted cash flow model with analysts' growth forecasts. The risk premium $rp = k - i$. The average of analysts' growth forecasts is g. *Div. yield* is expected dividend per share divided by price per share.

Year	Div. Yield	g	k	i	$rp = k - i$
1982	6.89	12.73	19.62	12.76	6.86
1983	5.24	12.60	17.86	11.18	6.67
1984	5.55	12.02	17.57	12.39	5.18
1985	4.97	11.45	16.42	10.79	5.63
1986	4.08	11.05	15.13	7.80	7.34
1987	3.64	11.01	14.65	8.58	6.07
1988	4.27	11.00	15.27	8.96	6.31
1989	3.95	11.08	15.03	8.45	6.58
1990	4.03	11.69	15.72	8.61	7.11
1991	3.64	11.99	15.63	8.14	7.50
1992	3.35	12.13	15.47	7.67	7.81
1993	3.15	11.63	14.78	6.60	8.18
1994	3.19	11.47	14.66	7.37	7.29
1995	3.04	11.51	14.55	6.88	7.67
1996	2.60	11.89	14.49	6.70	7.79
1997	2.18	12.60	14.78	6.60	8.17
1998	1.80	12.95	14.75	5.58	9.17
Average	3.86	11.81	15.67	8.53	7.14

government bonds is 7.14%, slightly higher than the 6.47% average for 1982 to 1991 reported by Harris and Marston (1992). For comparison purposes, Exhibit 3 contains historical returns and risk premia. The average expectational risk premium reported in Exhibit 2 is approximately equal to the arithmetic (7.5%) long-term differential between returns on stocks and long-term government bonds.[6]

[6]Interestingly, for the 1982-1996 period the arithmetic spread between large company stocks and long-term government bonds was only 3.3% per year. The downward trend in interest rates resulted in average annual returns of 14.1% on long-term government bonds over this horizon. Some (e.g., Ibbotson, 1997) argue that only the income (not total) return on bonds should be subtracted in calculating risk premia.

Exhibit 2 shows the estimated risk premium changes over time, suggesting changes in the market's perception of the incremental risk of investing in equity rather than debt securities. Scanning the last column of Exhibit 2, the risk premium is higher in the 1990s than earlier and especially so in late 1997 and 1998. Our DCF results provide no evidence to support the notion of a declining risk premium in the 1990s as a driver of the strong run up in equity prices.

A striking feature in Exhibit 2 is the relative stability of the estimates of k. After dropping (along with interest rates) in the early and mid-1980s, the average annual value of k has remained within a 75 basis point range around 15% for over a decade. Moreover, this stability arises despite some variability in the

Exhibit 3. Average Historical Returns on Bonds, Stocks, Bills, and Inflation in the US, 1926-1998

Historical Return Realizations	Geometric Mean	Arithmetic Mean
Common Stock (Large Company)	11.2%	13.2%
Long-term Government Bonds	5.3	5.7
Treasury Bills	3.8	3.8
Inflation Rate	3.1	3.2

Source: Ibbotson Associates, Inc. *1999 Stocks, Bonds, Bills and Inflation*, 1999 Yearbook.

underlying dividend yield and growth components of k as Exhibit 2 illustrates. The results suggest that k is more stable than government interest rates. Such relative stability of k translates into parallel changes in the market risk premium. In a subsequent section, we examine whether changes in our market risk premium estimates appear linked to interest rate conditions and a number of proxies for risk.

We explored the sensitivity of the results to our screening procedures in selecting companies. The reported results screen out all non-dividend paying stocks on the premise that use of the DCF model is inappropriate in such cases. The dividend screen eliminates an average of 55 companies per month. In a given month, we also screen out firms with fewer than three analysts' forecasts, or if the standard deviation around the mean forecast exceeds 20%. When the analysis is repeated without any of the three screens, the average risk premium over the sample period increased by only 40 basis points, from 7.14% to 7.54%. The beta of the sample firms also was estimated and the sample average was one, suggesting that the screens do not systematically remove low or high-risk firms. (Specifically, using firms in the screened sample as of December 1997 (the last date for which we had CRSP return data), we used ordinary least squares regressions to estimate beta for each stock using the prior 60 months of data and the CRSP return (SPRTRN) as the market index. The value-weighted average of the individual betas was 1.00.)

The results reported here use firms in the S&P500 as reported by COMPUSTAT in September 1998. This could create a survivorship bias, especially in the earlier months of the sample. We compared our current results to those obtained in Harris and Marston (1992) for which there was data to update the S&P500 composition each month. For the overlapping period, January 1982-May 1991, the two procedures yield the same average market risk premium, 6.47%. This suggests that the firms departing from or entering the S&P500 index do so for a number of reasons with no discernable effect on the overall estimated S&P500 market risk premium.

IV. Changes in the Market Risk Premium Over Time

With changes in the economy and financial markets, equity investments may be perceived to change in risk. For instance, investor sentiment about future business conditions likely affects attitudes about the riskiness of equity investments compared to investments in the bond markets. Moreover, since bonds are risky investments themselves, equity risk premia (relative to bonds) could change due to changes in perceived riskiness of bonds, even if equities displayed no shifts in risk.

In earlier work covering the 1982-1991 period, Harris and Marston (1992) reported regression results indicating that the market premium decreased with the level of government interest rates and increased with the spread between corporate and government bond yields (BSPREAD). This bond yield spread was interpreted as a time series proxy for equity risk. In this paper, we introduce three additional *ex ante* measures of risk shown in Exhibit 1: CON, DISP, and VOL. The three measures come from three independent sets of data and are supplied by different agents in the economy (consumers, equity analysts, and investors (via option and share price data)). Exhibit 4 provides summary data on all four of these risk measures.

Exhibit 5 replicates and updates earlier analysis by Harris and Marston (1992). The results confirm the earlier patterns. For the entire sample period, Panel A shows that risk premia are negatively related to interest rates. This negative relationship is also true for both

OLS regressions with levels of variables generally showed severe autocorrelation. As a result, we used the Prais-Winsten method (on levels of variables) and also OLS regressions on first differences of variables. Since both methods yielded similar results and the latter had more stable coefficients across specifications, we report only the results using first differences. Tests using Durbin-Watson statistics from regressions in Exhibits 5 and 6 do not accept the hypothesis of autocorrelated errors (tests at .01 significance level, see Johnston 1984). We also estimated the first difference model without an intercept and obtained estimates almost identical to those reported.

Exhibit 4. Descriptive Statistics on *Ex Ante* Risk Measures

Entries are based on monthly data. BSPREAD is the spread between yields on long-term corporate and government bonds. CON is the consumer confidence index. DISP measures the dispersion of analysts' forecasts of earnings growth. VOL is the volatility on the S&P500 index implied by options data. Variables are expressed in decimal form, (e.g., 12% = .12)

Panel A. Variables are Monthly Levels

	Mean	Standard Deviation	Minimum	Maximum
BSPREAD	0123	0040	0070	0254
CON	9504	2242	473	1 382
DISP	0349	0070	0285	0687
VOL	1599	0697	0765	6085

Panel B. Variables are Monthly Changes

	Mean	Standard Deviation	Minimum	Maximum
BSPREAD	00001	0011	- 0034	0036
CON	0030	0549	- 2300	2170
DISP	00002	0024	- 0160	0154
VOL	0008	0592	- 2156	4081

Panel C. Correlation Coefficients for Monthly Changes

	BSPREAD	CON	DISP	VOL
BSPREAD	1 00	- 16**	054	22*
CON	16**	1 00	065	- 09
DISP	054	065	1 00	027
VOL	22	- 09	027	1 00

**Significantly different from zero at the 05 level
*Significantly different from zero at the 01 level

the 1980s and 1990s as displayed in Panels B and C. For the entire 1982 to 1998 period, the addition of the yield spread risk proxy to the regressions lowers the magnitude of the coefficient on government bond yields, as can be seen by comparing Equations (1) and (2) of Panel A. Furthermore, the coefficient of the yield spread (0 488) is itself significantly positive. This pattern suggests that a reduction in the risk differential between investment in government bonds and in corporate bonds is translated into a lower equity market risk premium.

In major respects, the results in Exhibit 5 parallel earlier findings. The market risk premium changes over time and appears inversely related to government interest rates but is positively related to the bond yield spread, which proxies for the incremental risk of investing in equities as opposed to government bonds. One striking feature is the large negative coefficients on government bond yields. The coefficients indicate the equity risk premium declines by over 70 basis points for a 100 basis point increase in government interest rates.[5] This inverse relationship suggests

[5]The Exhibit 5 coefficients on r are significantly different from -1.0 suggesting that equity required returns do respond to interest rate changes. However, the large negative coefficients imply only minor adjustments of required returns to interest rate changes since the risk premium declines. In earlier work (Harris and Marston, 1992) the coefficient was significantly negative but not as large in absolute value. In that earlier work, we reported results using the Prais-Winsten estimators. When we use that estimation technique and recreate the second regression in Exhibit 5, the coefficient for r is - 584 (t = - 12.23) for the entire sample period 1982-1998

Exhibit 5. Changes in the Market Equity Risk Premium Over Time

The exhibit reports regression coefficients (t-values). Regression estimates use all variables expressed as monthly changes to correct for autocorrelation. The dependent variable is the market equity risk premium for the S&P500 index. BSPREAD is the spread between yields on long-term corporate and government bonds. The yield to maturity on long-term government bonds is denoted as i. For purposes of the regression, variables are expressed in decimal form, (e.g., 12% = .12)

Time Period	Intercept	i	BSPREAD	R^2
A. 1982-1998	-.0002	-.869		.57
	(-1.49)	(-16.54)		
	.0002	.749	.488	.59
	(-1.11)	(-11.37)	(2.94)	
B. 1980s	-.0005	-.887		.56
	(-1.62)	(10.97)		
	-.0004	.759	.508	.57
	(-1.24)	(-7.42)	(1.99)	
C. 1990s	-.0000	.840		.64
	(-0.09)	(13.78)		
	-.0000	.757	.347	.65
	(0.01)	(-9.85)	(1.76)	

Exhibit 6. Changes in the Market Equity Risk Premium Over Time and Selected Measures of Risk

The exhibit reports regression coefficients (t-values). Regression estimates use all variables expressed as monthly changes to correct for autocorrelation. The dependent variable is the market equity risk premium for the S&P500 index. BSPREAD is the spread between yields on long-term corporate and government bonds. The yield to maturity on long-term government bonds is denoted as i. CON is the consumer confidence index. DISP measures the dispersion of analysts' forecasts of earnings growth. VOL is the volatility on the S&P500 index implied by options data. For purposes of the regression, variables are expressed in decimal form, (e.g., 12% = .12)

Time Period		Intercept	i	BSPREAD	CON	DISP	VOL	Adj. R^2
A. 1982-1998								
	(1)	0.0002			-0.014			0.05
		(.97)			(-3.50)			
	(2)	-0.0001	-0.737	0.453	-0.007			0.60
		(-.96)	(-11.31)	(2.76)	(-2.48)			
	(3)	0.0002				0.224		0.02
		(.79)				(2.38)		
	(4)	0.0001	-0.733	0.433	-0.007	0.185		0.62
		(-.93)	(-11.49)	(2.69)	(-2.77)	(3.13)		
B. May 1986-1998	(5)	0.0000	0.818	0.420	-0.005	0.378		0.68
		(.06)	(11.21)	(2.52)	(-2.23)	(3.77)		
	(6)	0.0001					0.011	0.05
		(.53)					(2.89)	
	(7)	0.0000	-0.831	0.326	-0.005	0.372	0.006	0.69
		(.02)	(-11.52)	(1.95)	(-2.12)	(3.77)	(2.66)	

much greater stability in equity required returns than is often assumed. For instance, standard application of the CAPM suggests a one-to-one change in equity returns and government bond yields.

Exhibit 6 introduces three additional proxies for risk and explores whether these variables, either individually or collectively, are correlated with the market premium. Since the estimates of implied volatility start in May 1986, the exhibit shows results for both the entire sample period and for the period during which we can introduce all variables. Entered individually each of the three variables is significantly linked to the risk premium with the coefficient having the expected sign. For instance, in regression (1) the coefficient on CON is -.014, which is significantly different from zero (t = -3.50). The negative coefficient signals that higher consumer confidence is linked to a lower market premium. The positive coefficients on VOL and DISP indicate the equity risk premium increases with both market volatility and disagreement among analysts. The effects of the three variables appear largely unaffected by adding other variables. For instance, in regression (4) the coefficients on CON and DISP both remain significant and are similar in magnitude to the coefficients in single variable regressions.[9]

Even in the presence of the new risk variables, Exhibit 6 shows that the market risk premium is affected by interest rate conditions. The large negative coefficient on government bond rates implies large reductions in the equity premium as interest rates rise. One feature of our data may contribute to the observed negative relationship between the market risk premium and the level of interest rates. Specifically, if analysts are slow to report updates in their growth forecasts, changes in the estimated k would not adjust fully with changes in the interest rate even if the true risk premium were constant. To address the impact of "stickiness" in the measurement of k, we formed "quarterly" measures of the risk premium that treat k as an average over the quarter. Specifically, we take the value of k at the end of a quarter and subtract from it the average value of i for the months ending when k is measured. For instance, to form the risk premium for March 1998, the average value of i for January, February, and March is subtracted from the March value of k. This approach assumes that, in March, k still reflects values of g that have not been updated from the prior two months. The quarterly measure of risk premium then is paired with the average values of the other variables for the quarter. For instance, the March 1998 "quarterly" risk premium would be paired with averaged values of BSPREAD over the January through March period. To avoid overlapping observations for the independent variables, we use only every third month (March, June, September, December) in the sample.

As reported in Exhibit 7, sensitivity analysis using "quarterly" observations suggests that delays in updating may be responsible for a portion, but not all, of the observed negative relationship between the market premium and interest rates. For example, when quarterly observations are used, the coefficient on i in regression (2) of Exhibit 7 is -.527, well below the earlier estimates but still significantly negative.[10]

As an additional test, movements in the bond risk premium (BSPREAD) are examined. Since BSPREAD is constructed directly from bond yield data, it does not have the potential for reporting lags that may affect analysts' growth forecasts. Regression 3 in Exhibit 7 shows BSPREAD is negatively linked to government rates and significantly so. While the equity premium need not move in the same pattern as the corporate bond premium, the negative coefficient on BSPREAD suggests that our earlier results are not due solely to "stickiness" in measurements of market required returns.

The results in Exhibit 7 suggest that the inverse relationship between interest rates and the market risk premium may not be as pronounced as suggested in earlier exhibits. Still, there appears to be a significant negative link between the equity risk premium and government interest rates. The quarterly results in Exhibit 7 would suggest about a 50 basis point change in risk premium for each 100 basis point movement in interest rates.

Overall, the *ex ante* estimates of the market risk premium are significantly linked to *ex ante* proxies for risk. Such a link suggests that investors modify their required returns in response to perceived changes in the environment. The findings provide some comfort that our risk premium estimates are capturing, at least

[9] Realized equity returns are difficult to predict out of sample (see Goyal and Welch, 1999). Our approach is different in that we look at expectational risk premia which are much more stable. For instance, when we estimate regression coefficients (using the specification shown in regression 7 of Exhibit 6) and apply them out of sample we obtain "predictions" of expectational risk premia that are significantly more accurate (better than the .01 level) than a no change forecast. We use a "rolling regression" approach using data through December 1991 to get coefficients to predict the risk premium in January 1992. We repeat the procedure moving forward a month and dropping the oldest month of data from the regression. Details are available from the authors.

[10] Sensitivity analysis for the 1982-1989 and 1990-1998 subperiods yields results similar to those reported.

We thank Bob Conroy for suggesting use of BSPREAD.

Regression 3 in Exhibit 7 appears to have autocorrelated errors; the Durbin-Watson (DW) statistic rejects the hypothesis of no autocorrelation. However, in subperiod analysis the DW statistic for the 1990-98 period is consistent with no autocorrelation and the coefficient on i is essentially the same (-.24, t = -8.05) as reported in Exhibit 7.

10899

Exhibit 7. Regressions Using Alternate Measures of Risk Premia to Analyze Potential Effects of Reporting Lags in Analysts' Forecasts

The exhibit reports regression coefficients (*t*-values). Regression estimates use all variables expressed as changes (monthly or quarterly) to correct for autocorrelation. BSPREAD is the spread between yields on long-term corporate and government bonds. *rp* is the risk premium on the S&P500 index. The yield to maturity on long-term government bonds is denoted as *i*. For purposes of the regression, variables are expressed in decimal form. (e g , 12% 12)

Dependent Variable	Intercept	*i*	BSPREAD	*Adj.* R^2
(1) Equity Risk Premium (*rp*) Monthly Observations (same as Table V)	0002 (1 11)	- 749 (-11 37)	488 (2 94)	59
(2) Equity Risk Premium (*rp*) "Quarterly" nonoverlapping observations to account for lags in analyst reporting	- 0002 (- 49)	- 527 (-6 18)	550 (2 20)	60
(3) Corporate Bond Spread (BSPREAD) Monthly Observations	0001 (-1 90)	- 247 (-11.29)		38

in part, underlying changes in the economic environment. Moreover, each of the risk measures appears to contain relevant information for investors. The market risk premium is negatively related to the level of consumer confidence and positively linked to interest rate spreads between corporate and government debt, disagreement among analysts in their forecasts of earnings growth, and the implied volatility of equity returns as revealed in options data.

V. Conclusions

Shareholder required rates of return and risk premia should be based on theories about investors' expectations for the future. In practice, however, risk premia are typically estimated using averages of historical returns. This paper applies an alternate approach to estimating risk premia that employs publicly available expectational data. The resultant average market equity risk premium over government bonds is comparable in magnitude to long-term differences (1926 to 1998) in historical returns between stocks and bonds. As a result, our evidence does not resolve the equity premium puzzle, rather, the results suggest investors still expect to receive large spreads to invest in equity versus debt instruments.

There is strong evidence, however, that the market risk premium changes over time. Moreover, these changes appear linked to the level of interest rates as well as *ex ante* proxies for risk drawn from interest rate spreads in the bond market, consumer confidence in future economic conditions, disagreement among financial analysts in their forecasts and the volatility of equity returns implied by options data. The significant economic links between the market premium and a wide array of risk variables suggests that the notion of a constant risk premium over time is not an adequate explanation of pricing in equity versus debt markets.

These results have implications for practice. First, at least on average, the estimates suggest a market premium roughly comparable to long-term historical spreads in returns between stocks and bonds. Our conjecture is that, if anything, the estimates are on the high side and thus establish an upper bound on the market premium. Second, the results suggest that use of a constant risk premium will not fully capture changes in investor return requirements. As a specific example, our findings indicate that common application of models such as the CAPM will overstate changes in shareholder return requirements when government interest rates change. Rather than a one-for-one change with interest rates implied by use of constant risk premium, the results indicate that equity required returns for average risk stocks likely change by half (or less) of the change in interest rates. However, the picture is considerably more complicated as shown by the linkages between the risk premium and other attributes of risk.

Ultimately, our research does not resolve the answer to the question "What is the right market risk premium?" Perhaps more importantly, our work suggests that the answer is conditional on a number of features in the economy not an absolute. We hope that future research will harness *ex ante* data to provide additional guidance to best practice in using a market premium to improve financial decisions. ■

References

Asness, C. S., 2000, "Stocks versus Bonds: Explaining the Equity Risk Premium," *Financial Analysts Journal* 56 (No. 2, March/April), 96-113.

Boebel, R. B., 1991, "The Information Content of Financial Analysts' Forecasts: Short-term vs. Long-term," University of North Carolina at Chapel Hill Dissertation. University Microfilms International order number 9207936.

Boebel, R. B., M.N. Gultekin and R. S. Harris, 1993, "Financial Analysts' Forecast of Corporate Earnings Growth: Top Down Versus Bottom Up," in *Handbook of Security Analyst Forecasting and Asset Allocation*, John B. Guerard, Jr. and Mustafa N. Gultekin (eds.), London, JAI Press, 185-192.

Brigham, D., D. Shome, and S. Vinson, 1985, "The Risk Premium Approach to Measuring A Utility's Cost of Equity," *Financial Management* 14 (No. 1, Spring), 33-45.

Brown, L. D., 1993, "Earnings Forecasting Research: Its Implications for Capital Market Research," *International Journal of Forecasting* 9 (No. 3, November), 295-320.

Brown, L. D., 1997, "Analyst Forecasting Errors: Additional Evidence," *Financial Analysts Journal* 53 (No. 6, November/December), 81-90.

Bruner, R. F., K. Eades, R. Harris and R. Higgins, 1998, "Best Practices in Estimating the Cost of Capital: Survey and Synthesis," *Financial Practice and Education* 8 (No. 1, Spring/Summer), 13-28.

Carleton, W. T. and J. Lakonishok, 1985, "Risk and Return on Equity: The Use and Misuse of Historical Estimates," *Financial Analysts Journal* 41 (No. 1, January/February), 38-47.

Chan, L., Y. Hamao, and J. Lakonishok, 1991, "Fundamental and Stock Returns in Japan," *Journal of Finance* 46 (No. 5, December) 1739-64.

Cragg, J. and B.G. Malkiel, 1982, *Expectations and the Structure of Share Prices*, National Bureau of Economic Research, Chicago, University of Chicago Press.

Goldman Sachs, 1999, *EAI Company Analysis*, (January).

Goyal, A. and I. Welch, 1999, "Predicting the Equity Premium," Anderson Graduate School of Management at UCLA (August 14).

Harris, R., 1986, "Using Analysts' Growth Forecasts to Estimate Shareholder Required Rates of Return," *Financial Management* 15 (No. 1, Spring), 58-67.

Harris, R. and F. Marston, 1992, "Estimating Shareholder Risk Premia Using Analysts' Growth Forecasts," *Financial Management* 21 (No. 2, Summer), 63-70.

Ibbotson Associates, Inc., *1998 Cost of Capital Quarterly*, 1998 Yearbook.

Ibbotson Associates, Inc., *1997 Stocks, Bonds, Bills, and Inflation*, 1997 Yearbook.

Johnston, J., 1984, *Econometric Methods*, New York, McGraw-Hill. 3rd edition.

Kennedy, P., 1985, *A Guide to Econometrics*, Cambridge, MIT Press. 2nd edition.

Lamdin, D., 2001, "Estimating the Cost of Equity for Corporations that Repurchase: Theory and Application,"*Engineering Economist*, 46 (No. 1), 53-63.

Madalla, G.S., 1977, *Econometrics*, New York, McGraw-Hill.

Malkiel, B., 1979, "The Capital Formation Problem in the United States," *Journal of Finance* 34 (No. 2, May), 291-306.

Malkiel, B., 1982, "Risk and Return: A New Look," in *The Changing Role of Debt and Equity in Financing US Capital Formation*, B.B. Friedman (ed.), National Bureau of Economic Research, Chicago, University of Chicago Press.

Marston, F., R. Harris, and P. Crawford, 1992, "Risk and Return in Equity Markets: Evidence Using Financial Analysts' Forecasts," in *Handbook of Security Analysts Forecasting and Asset Allocation*, J. Guerard and M. Gultekin (eds.), Greenwich, CT, JAI Press, 101-122.

Marston, F. and R. Harris, 1993, "Risk, Return and Equilibrium: A Revisit Using Expected Returns," *Financial Review*, 28 (No. 1, February).

Siegel, J., 1999, "The Shrinking Equity Premium," *The Journal of Portfolio Management* (Fall), 10-17.

Siegel, J., and R. Thaler, 1997, "Anomalies: The Equity Premium Puzzle," *Journal of Economic Perspectives*, 11 (No. 1, Winter), 191-200.

VanderWeide, J. and W.T. Carleton, 1988, "Investor Growth Expectations: Analysts vs. History," *Journal of Portfolio Management* (Spring), 78-82.

Welch, I., 2000, "Views of Financial Economists on the Equity Premium and on Professional Controversies," *Journal of Business* 73 (No. 4, October) 501-537.

Cost of Capital Estimation

The Risk Premium Approach to Measuring a Utility's Cost of Equity

Eugene F. Brigham, Dilip K. Shome, and Steve R. Vinson

Eugene F. Brigham and Dilip K. Shome are faculty members of the University of Florida and the Virginia Polytechnic Institute and State University, respectively; Steve R. Vinson is affiliated with AT&T Communications

■ In the mid-1960s, Myron Gordon and others began applying the theory of finance to help estimate utilities' costs of capital. Previously, the standard approach in cost of equity studies was the "comparable earnings method," which involved selecting a sample of unregulated companies whose investment risk was judged to be comparable to that of the utility in question, calculating the average return on book equity (ROE) of these sample companies, and setting the utility's service rates at a level that would permit the utility to achieve the same ROE as comparable companies. This procedure has now been thoroughly discredited (see Robichek [15]), and it has been replaced by three market-oriented (as opposed to accounting-oriented) approaches: (i) the DCF method, (ii) the bond-yield-plus-risk-premium method, and (iii) the CAPM, which is a specific version of the generalized bond-yield-plus-risk-premium approach.

Our purpose in this paper is to discuss the risk-premium approach, including the market risk premium that is used in the CAPM. First, we critique the various procedures that have been used in the past to estimate risk premiums. Second, we present some data on estimated risk premiums since 1965. Third, we examine the relationship between equity risk premiums and the level of interest rates, because it is important, for purposes of estimating the cost of capital, to know just how stable the relationship between risk premiums and interest rates is over time. If stability exists, then one can estimate the cost of equity at any point in time as a function of interest rates as reported in *The Wall Street Journal*, the *Federal Reserve Bulletin*, or some similar source.[1] Fourth, while we do not discuss the CAPM directly, our analysis does have some important implications for selecting a market risk premium for use in that model. Our focus is on utilities, but the methodology is applicable to the estimation of the cost of

[1]For example, the Federal Energy Regulatory Commission's Staff recently proposed that a risk premium be estimated every two years and that, between estimation dates, the last-determined risk premium be added to the current yield on ten-year Treasury bonds to obtain an estimate of the cost of equity to an average utility (Docket RM 80-36). Subsequently, the FCC made a similar proposal ("Notice of Proposed Rulemaking," August 13, 1984, Docket No 84-800). Obviously, the validity of such procedures depends on (i) the accuracy of the risk premium estimate and (ii) the stability of the relationship between risk premiums and interest rates. Both proposals are still under review.

equity for any publicly traded firm, and also for non-traded firms for which an appropriate risk class can be assessed, including divisions of publicly traded corporations.[2]

Alternative Procedures for Estimating Risk Premiums

In a review of both rate cases and the academic literature, we have identified three basic methods for estimating equity risk premiums: (i) the *ex post*, or historic, yield spread method; (ii) the survey method; and (iii) an *ex ante* yield spread method based on DCF analysis.[3] In this section, we briefly review these three methods.

Historic Risk Premiums

A number of researchers, most notably Ibbotson and Sinquefield [12], have calculated historic holding period returns on different securities and then estimated risk premiums as follows:

$$\text{Historic Risk Premium} = \begin{pmatrix}\text{Average of the} \\ \text{annual returns on} \\ \text{a stock index for} \\ \text{a particular} \\ \text{past period}\end{pmatrix} - \begin{pmatrix}\text{Average of the} \\ \text{annual returns on} \\ \text{a bond index for} \\ \text{the same} \\ \text{past period}\end{pmatrix}. \quad (1)$$

Ibbotson and Sinquefield (I&S) calculated both arithmetic and geometric average returns, but most of their risk-premium discussion was in terms of the geometric averages. Also, they used both corporate and Treasury bond indices, as well as a T-bill index, and they analyzed all possible holding periods since 1926. The I&S study has been employed in numerous rate cases in two ways: (i) directly, where the I&S historic risk premium is added to a company's bond yield to obtain an estimate of its cost of equity, and (ii) indirectly, where I&S data are used to estimate the market risk premium in CAPM studies.

There are both conceptual and measurement problems with using I&S data for purposes of estimating the cost of capital. Conceptually, there is no compelling reason to think that investors expect the same relative returns that were earned in the past. Indeed, evidence presented in the following sections indicates that relative expected returns should, and do, vary significantly over time. Empirically, the measured historic premium is sensitive both to the choice of estimation horizon and to the end points. These choices are essentially arbitrary, yet they can result in significant differences in the final outcome. These measurement problems are common to most forecasts based on time series data.

The Survey Approach

One obvious way to estimate equity risk premiums is to poll investors. Charles Benore [1], the senior utility analyst for Paine Webber Mitchell Hutchins, a leading institutional brokerage house, conducts such a survey of major institutional investors annually. His 1983 results are reported in Exhibit 1.

Exhibit 1. Results of Risk Premium Survey, 1983*

Assuming a double A, long-term utility bond currently yields 12½%, the common stock for the same company would be fairly priced relative to the bond if its expected return was as follows:

Total Return	Indicated Risk Premium (basis points)	Percent of Respondents
over 20½%	over 800	
20½%	800	
19½%	700	
18½%	600	10%
17½%	500	8%
16½%	400	29%
15½%	300	35%
14½%	200	16%
13½%	100	0%
under 13½%	under 100	1%
Weighted average	358	100%

*Benore's questionnaire included the first two columns, while his third column provided a space for the respondents to indicate which risk premium they thought applied. We summarized Benore's responses in the frequency distribution given in Column 3. Also, in his questionnaire each year, Benore adjusts the double A bond yield and the total return (Column 1) to reflect current market conditions. Both the question above and the responses to it were taken from the survey conducted in April 1983.

[2]The FCC is particularly interested in risk-premium methodologies, because (i) only eighteen of the 1,400 telephone companies it regulates have publicly-traded stock, and hence offer the possibility of DCF analysis, and (ii) most of the publicly-traded telephone companies have both regulated and unregulated assets, so a corporate DCF cost might not be applicable to the regulated units of the companies.

[3]In rate cases, some witnesses also have calculated the differential between the yield to maturity (YTM) of a company's bonds and its concurrent ROE, and then called this differential a risk premium. In general, this procedure is unsound, because the YTM on a bond is a *future expected* return on the bond's *market value*, while the ROE is the *past realized* return on the stock's *book value*. Thus, comparing YTMs and ROEs is like comparing apples and oranges.

Benore's results, as measured by the average risk premiums, have varied over the years as follows:

Year	Average RP (basis points)
1978	491
1979	475
1980	423
1981	349
1982	275
1983	358

The survey approach is conceptually sound in that it attempts to measure investors' expectations regarding risk premiums, and the Benore data also seem to be carefully collected and processed. Therefore, the Benore studies do provide one useful basis for estimating risk premiums. However, as with most survey results, the possibility of biased responses and/or biased sampling always exists. For example, if the responding institutions are owners of utility stocks (and many of them are), and if the respondents think that the survey results might be used in a rate case, then they might bias upward their responses to help utilities obtain higher authorized returns. Also, Benore surveys large institutional investors, whereas a high percentage of utility stocks are owned by individuals rather than institutions, so there is a question as to whether his reported risk premiums are really based on the expectations of the "representative" investor. Finally, from a pragmatic standpoint, there is a question as to how to use the Benore data for utilities that are not rated AA. The Benore premiums can be applied as an add-on to the own-company bond yields of any given utility only if it can be assumed that the premiums are constant across bond rating classes. *A priori*, there is no reason to believe that the premiums will be constant.

DCF-Based *Ex Ante* Risk Premiums

In a number of studies, the DCF model has been used to estimate the *ex ante* market risk premium, RP_M. Here, one estimates the average expected future return on equity for a group of stocks, k_M, and then subtracts the concurrent risk-free rate, R_F, as proxied by the yield to maturity on either corporate or Treasury securities:[4]

$$RP_M = k_M - R_F. \tag{2}$$

Conceptually, this procedure is exactly like the I&S approach except that one makes direct estimates of future expected returns on stocks and bonds rather than assuming that investors expect future returns to mirror past returns.

The most difficult task, of course, is to obtain a valid estimate of k_M, the expected rate of return on the market. Several studies have attempted to estimate DCF risk premiums for the utility industry and for other stock market indices. Two of these are summarized next.

Vandell and Kester. In a recently published monograph, Vandell and Kester [18] estimated *ex ante* risk premiums for the period from 1944 to 1978. R_F was measured both by the yield on 90-day T-bills and by the yield on the Standard and Poor's AA Utility Bond Index. They measured k_M as the average expected return on the S&P's 500 Index, with the expected return on individual securities estimated as follows:

$$k_i = \left(\frac{D_1}{P_0}\right)_i + g_i, \tag{3}$$

where,

D_1 = dividend per share expected over the next twelve months,
P_0 = current stock price,
g = estimated long-term constant growth rate, and
i = the i^{th} stock.

To estimate g_i, Vandell and Kester developed fifteen forecasting models based on both exponential smoothing and trend-line forecasts of earnings and dividends, and they used historic data over several estimating horizons. Vandell and Kester themselves acknowledge that, like the Ibbotson-Sinquefield premiums, their analysis is subject to potential errors associated with trying to estimate expected future growth purely from past data. We shall have more to say about this point later.

---

[4]In this analysis, most people have used yields on long-term bonds rather than short-term money market instruments. It is recognized that long-term bonds, even Treasury bonds, are not risk free, so an RP_M based on these debt instruments is smaller than it would be if there were some better proxy to the long-term riskless rate. People have attempted to use the T-bill rate for R_F, but the T-bill rate embodies a different average inflation premium than stocks, and it is subject to random fluctuations caused by monetary policy, international currency flows, and other factors. Thus, many people believe that for cost of capital purposes, R_F should be based on long-term securities.

We did test to see how debt maturities would affect our calculated risk premiums. If a short-term rate such as the 30-day T-bill rate is used, measured risk premiums jump around widely and, so far as we could tell, randomly. The choice of a maturity in the 10- to 30-year range has little effect, as the yield curve is generally fairly flat in that range.

Malkiel. Malkiel [14] estimated equity risk premiums for the Dow Jones Industrials using the DCF model. Recognizing that the constant dividend growth assumption may not be valid, Malkiel used a nonconstant version of the DCF model. Also, rather than rely exclusively on historic data, he based his growth rates on Value Line's five-year earnings growth forecasts plus the assumption that each company's growth rate would, after an initial five-year period, move toward a long-run real national growth rate of four percent. He also used ten-year maturity government bonds as a proxy for the riskless rate. Malkiel reported that he tested the sensitivity of his results against a number of different types of growth rates, but, in his words, "The results are remarkably robust, and the estimated risk premiums are all very similar." Malkiel's is, to the best of our knowledge, the first risk-premium study that uses analysts' forecasts. A discussion of analysts' forecasts follows.

Security Analysts' Growth Forecasts

Ex ante DCF risk premium estimates can be based either on expected growth rates developed from time series data, such as Vandell and Kester used, or on analysts' forecasts, such as Malkiel used. Although there is nothing inherently wrong with time series-based growth rates, an increasing body of evidence suggests that primary reliance should be placed on analysts' growth rates. First, we note that the observed market price of a stock reflects the consensus view of investors regarding its future growth. Second, we know that most large brokerage houses, the larger institutional investors, and many investment advisory organizations employ security analysts who forecast future EPS and DPS, and, to the extent that investors rely on analysts' forecasts, the consensus of analysts' forecasts is embodied in market prices. Third, there have been literally dozens of academic research papers dealing with the accuracy of analysts' forecasts, as well as with the extent to which investors actually use them. For example, Cragg and Malkiel [7] and Brown and Rozeff [5] determined that security analysts' forecasts are more relevant in valuing common stocks and estimating the cost of capital than are forecasts based solely on historic time series. Stanley, Lewellen, and Schlarbaum [16] and Linke [13] investigated the importance of analysts' forecasts and recommendations to the investment decisions of individual and institutional investors. Both studies indicate that investors rely heavily on analysts' reports and incorporate analysts' forecast information in the formation of their expectations about stock returns. A representative listing of other work supporting the use of analysts' forecasts is included in the References section. Thus, evidence in the current literature indicates that (i) analysts' forecasts are superior to forecasts based solely on time series data, and (ii) investors do rely on analysts' forecasts. Accordingly, we based our cost of equity, and hence risk premium estimates, on analysts' forecast data.[5]

Risk Premium Estimates

For purposes of estimating the cost of capital using the risk premium approach, it is necessary either that the risk premiums be time-invariant or that there exists a predictable relationship between risk premiums and interest rates. If the premiums are constant over time, then the constant premium could be added to the prevailing interest rate. Alternatively, if there exists a stable relationship between risk premiums and interest rates, it could be used to predict the risk premium from the prevailing interest rate

To test for stability, we obviously need to calculate risk premiums over a fairly long period of time. Prior to 1980, the only consistent set of data we could find came from Value Line, and, because of the work involved, we could develop risk premiums only once a year (on January 1). Beginning in 1980, however, we began collecting and analyzing Value Line data on a monthly basis, and in 1981 we added monthly estimates from Merrill Lynch and Salomon Brothers to our data base. Finally, in mid-1983, we expanded our analysis to include the IBES data.

Annual Data and Results, 1966–1984

Over the period 1966–1984, we used Value Line data to estimate risk premiums both for the electric utility industry and for industrial companies, using the companies included in the Dow Jones Industrial and Utility averages as representative of the two groups. Value Line makes a five-year growth rate forecast, but it also gives data from which one can develop a longer-term forecast. Since DCF theory calls for a truly long-term (infinite horizon) growth rate, we concluded that it was better to develop and use such a forecast than to

[5]Recently, a new type of service that summarizes the key data from most analysts' reports has become available. We are aware of two sources of such services, the Lynch, Jones, and Ryan's Institutional Brokers Estimate System (IBES) and Zack's Icarus Investment Service. IBES and the Icarus Service gather data from both buy-side and sell-side analysts and provide it to subscribers on a monthly basis in both a printed and a computer-readable format.

Exhibit 2. Estimated Annual Risk Premiums, Nonconstant (Value Line) Model, 1966–1984

January 1 of the Year Reported	Dow Jones Electrics			Dow Jones Industrials			
	k_{Avg}	R_F	RP	k_{Avg}	R_F	RP	(3) − (6)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1966	8 11%	4.50%	3 61%	9.56%	4 50%	5.06%	0.71
1967	9.00%	4 76%	4 24%	11.57%	4 76%	6 81%	0 62
1968	9 68%	5 59%	4 09%	10 56%	5.59%	4.97%	0 82
1969	9 34%	5.88%	3.46%	10.96%	5 88%	5.08%	0 68
1970	11.04%	6.91%	4 13%	12 22%	6.91%	5.31%	0.78
1971	10 80%	6.28%	4.52%	11.23%	6.28%	4.95%	0 91
1972	10 53%	6.00%	4.53%	11.09%	6.00%	5.09%	0 89
1973	11.37%	5 96%	5 41%	11 47%	5 96%	5.51%	0 98
1974	13 85%	7 29%	6 56%	12 38%	7 29%	5 09%	1.29
1975	16.63%	7.91%	8 72%	14.83%	7.91%	6.92%	1 26
1976	13 97%	8.23%	5 74%	13.32%	8.23%	5 09%	1 13
1977	12 96%	7.30%	5 66%	13 63%	7.30%	6 33%	0 89
1978	13 42%	7.87%	5 55%	14 75%	7.87%	6.88%	0 81
1979	14 92%	8 99%	5.93%	15.50%	8 99%	6 51%	0 91
1980	16 39%	10 18%	6.21%	16 53%	10.18%	6.35%	0.98
1981	17 61%	11 99%	5 62%	17.37%	11 99%	5.38%	1.04
1982	17.70%	14.00%	3 70%	19 30%	14 00%	5.30%	0.70
1983	16.30%	10.66%	5 64%	16.53%	10 66%	5 87%	0 96
1984	16 03%	11 97%	4.06%	15 72%	11 97%	3.75%	1.08

use the five-year prediction.[6] Therefore, we obtained data as of January 1 from Value Line for each of the Dow Jones companies and then solved for k, the expected rate of return, in the following equation:

$$P_0 = \sum_{t=1}^{n} \frac{D_t}{(1 + k)^t} + \left(\frac{D_n(1 + g_n)}{k - g_n}\right)\left(\frac{1}{1 + k}\right)^n. \quad (4)$$

Equation (4) is the standard nonconstant growth DCF model; P_0 is the current stock price; D_t represents the forecasted dividends during the nonconstant growth period; n is the years of nonconstant growth; D_n is the first constant growth dividend; and g_n is the constant, long-run growth rate after year n. Value Line provides D_t values for t = 1 and t = 4, and we interpolated to obtain D_2 and D_3. Value Line also gives estimates for ROE and for the retention rate (b) in the terminal year, n, so we can forecast the long-term growth rate as g_n = b(ROE). With all the values in Equation (4) specified except k, we can solve for k, which is the DCF rate of return that would result if the Value Line forecasts were met, and, hence, the DCF rate of return implied in the Value Line forecast.[7]

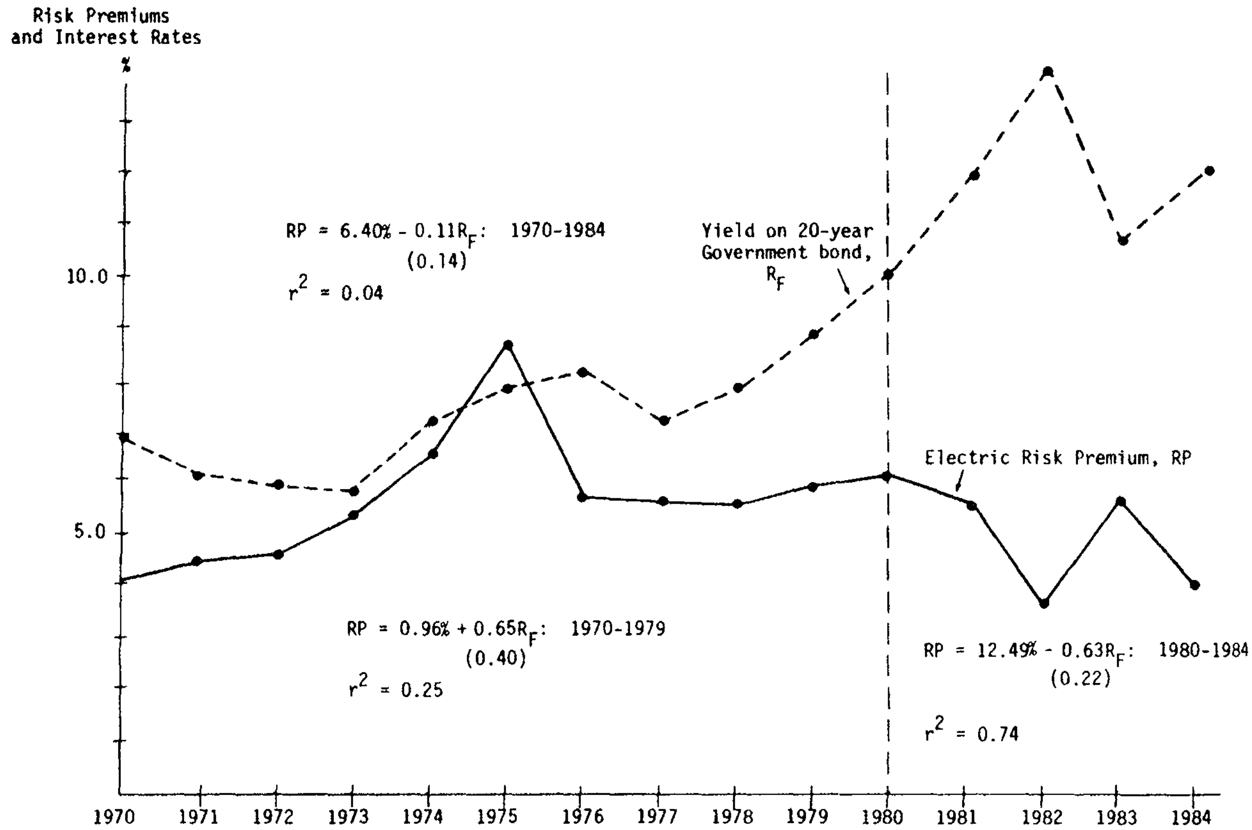
Having estimated a k value for each of the electric and industrial companies, we averaged them (using market-value weights) to obtain a k value for each group, after which we subtracted R_F (taken as the December 31 yield on twenty-year constant maturity Treasury bonds) to obtain the estimated risk premiums shown in Exhibit 2. The premiums for the electrics are plotted in Exhibit 3, along with interest rates. The following points are worthy of note:

1. Risk premiums fluctuate over time. As we shall see in the next section, fluctuations are even wider when measured on a monthly basis.
2. The last column of Exhibit 2 shows that risk premi-

[6] This is a debatable point. Cragg and Malkiel, as well as many practicing analysts, feel that most investors actually focus on five-year forecasts. Others, however, argue that five-year forecasts are too heavily influenced by base-year conditions and/or other nonpermanent conditions for use in the DCF model We note (i) that most published forecasts do indeed cover five years, (ii) that such forecasts are typically "normalized" in some fashion to alleviate the base-year problem, and iii) that for relatively stable companies like those in the Dow Jones averages, it generally does not matter greatly if one uses a normalized five-year or a longer-term forecast, because these companies meet the conditions of the constant-growth DCF model rather well

[7] Value Line actually makes an explicit price forecast for each stock, and one could use this price, along with the forecasted dividends, to develop an expected rate of return However, Value Line's forecasted stock price builds in a forecasted *change* in k Therefore, the forecasted price is inappropriate for use in estimating current values of k.

Exhibit 3. Equity Risk Premiums for Electric Utilities and Yields on 20-Year Government Bonds, 1970–1984*



*Standard errors of the coefficients are shown in parentheses below the coefficients

ums for the utilities increased relative to those for the industrials from the mid-1960s to the mid-1970s. Subsequently, the perceived riskiness of the two groups has, on average, been about the same.

3. Exhibit 3 shows that, from 1970 through 1979, utility risk premiums tended to have a positive association with interest rates: when interest rates rose, so did risk premiums, and vice versa. However, beginning in 1980, an inverse relationship appeared: rising interest rates led to declining risk premiums. We shall discuss this situation further in the next section.

Monthly Data and Results, 1980–1984

In early 1980, we began calculating risk premiums on a monthly basis. At that time, our only source of analysts' forecasts was Value Line, but beginning in 1981 we also obtained Merrill Lynch and Salomon Brothers' data, and then, in mid-1983, we obtained IBES data. Because our focus was on utilities, we restricted our monthly analysis to that group.

Our 1980–1984 monthly risk premium data, along with Treasury bond yields, are shown in Exhibits 4 and 5 and plotted in Exhibits 6, 7, and 8. Here are some comments on these Exhibits:

1. Risk premiums, like interest rates and stock prices, are volatile. Our data indicate that it would not be appropriate to estimate the cost of equity by adding the current cost of debt to a risk premium that had been estimated in the past. Current risk premiums should be matched with current interest rates.
2. Exhibit 6 confirms the 1980–1984 section of Exhibit 3 in that it shows a strong inverse relationship between interest rates and risk premiums; we shall discuss shortly why this relationship holds.
3. Exhibit 7 shows that while risk premiums based on Value Line, Merrill Lynch, and Salomon Brothers

Exhibit 4. Estimated Monthly Risk Premiums for Electric Utilities Using Analysts' Growth Forecasts, January 1980–June 1984

Beginning of Month	Value Line	Merrill Lynch	Salomon Brothers	Average Premiums	20-Year Treasury Bond Yield, Constant Maturity Series
Jan 1980	6.21%	NA	NA	6.21%	10.18%
Feb 1980	5.77%	NA	NA	5.77%	10.86%
Mar 1980	4.73%	NA	NA	4.73%	12.59%
Apr 1980	5.02%	NA	NA	5.02%	12.71%
May 1980	4.73%	NA	NA	4.73%	11.04%
Jun 1980	5.09%	NA	NA	5.09%	10.37%
Jul 1980	5.41%	NA	NA	5.41%	9.86%
Aug 1980	5.72%	NA	NA	5.72%	10.29%
Sep 1980	5.16%	NA	NA	5.16%	11.41%
Oct 1980	5.62%	NA	NA	5.62%	11.75%
Nov 1980	5.09%	NA	NA	5.09%	12.33%
Dec 1980	5.65%	NA	NA	5.65%	12.37%
Annual Avg	5.35%			5.35%	11.31%
Jan 1981	5.62%	4.76%	5.63%	5.34%	11.99%
Feb 1981	4.82%	4.87%	5.16%	4.95%	12.48%
Mar 1981	4.70%	3.73%	4.97%	4.47%	13.10%
Apr 1981	4.24%	3.23%	4.52%	4.00%	13.11%
May 1981	3.54%	3.24%	4.24%	3.67%	13.51%
Jun 1981	3.57%	4.04%	4.27%	3.96%	13.39%
Jul 1981	3.61%	3.63%	4.16%	3.80%	13.32%
Aug 1981	3.17%	3.05%	3.04%	3.09%	14.23%
Sep 1981	2.11%	2.24%	2.35%	2.23%	14.99%
Oct 1981	2.83%	2.64%	3.24%	2.90%	14.93%
Nov 1981	2.08%	2.49%	3.03%	2.53%	15.27%
Dec 1981	3.72%	3.45%	4.24%	3.80%	13.12%
Annual Avg	3.67%	3.45%	4.07%	3.73%	13.62%
Jan 1982	3.70%	3.37%	4.04%	3.70%	14.00%
Feb 1982	3.05%	3.37%	3.70%	3.37%	14.37%
Mar 1982	3.15%	3.28%	3.75%	3.39%	13.96%
Apr 1982	3.49%	3.61%	4.29%	3.80%	13.69%
May 1982	3.08%	4.25%	3.91%	3.75%	13.47%
Jun 1982	3.16%	4.51%	4.72%	4.13%	13.53%
Jul 1982	2.57%	4.21%	4.21%	3.66%	14.48%
Aug 1982	4.33%	4.83%	5.27%	4.81%	13.69%
Sep 1982	4.08%	5.14%	5.58%	4.93%	12.40%
Oct 1982	5.35%	5.24%	6.34%	5.64%	11.95%
Nov 1982	5.67%	5.95%	6.91%	6.18%	10.97%
Dec 1982	6.31%	6.71%	7.45%	6.82%	10.52%
Annual Avg	4.00%	4.54%	5.01%	4.52%	13.09%
Jan 1983	5.64%	6.04%	6.81%	6.16%	10.66%
Feb 1983	4.68%	5.99%	6.10%	5.59%	11.01%
Mar 1983	4.99%	6.89%	6.43%	6.10%	10.71%
Apr 1983	4.75%	5.82%	6.31%	5.63%	10.84%
May 1983	4.50%	6.41%	6.24%	5.72%	10.57%
Jun 1983	4.29%	5.21%	6.16%	5.22%	10.90%
Jul 1983	4.78%	5.72%	6.42%	5.64%	11.12%
Aug 1983	3.89%	4.74%	5.41%	4.68%	11.78%
Sep 1983	4.07%	4.90%	5.57%	4.85%	11.71%
Oct 1983	3.79%	4.64%	5.38%	4.60%	11.64%
Nov 1983	2.84%	3.77%	4.46%	3.69%	11.90%
Dec 1983	3.36%	4.27%	5.00%	4.21%	11.83%
Annual Avg	4.30%	5.37%	5.86%	5.17%	11.22%
Jan 1984	4.06%	5.04%	5.65%	4.92%	11.97%
Feb 1984	4.25%	5.37%	5.96%	5.19%	11.76%
Mar 1984	4.73%	6.05%	6.38%	5.72%	12.12%
Apr 1984	4.78%	5.33%	6.32%	5.48%	12.51%
May 1984	4.36%	5.30%	6.42%	5.36%	12.78%
Jun 1984	3.54%	4.00%	5.63%	4.39%	13.60%

Exhibit 5. Monthly Risk Premiums Based on IBES Data

Beginning of Month	Average of Merrill Lynch, Salomon Brothers, and Value Line Premiums for Dow Jones Electrics	IBES Premiums for Dow Jones Electrics	IBES Premiums for Entire Electric Industry
Aug 1983	4.68%	4.10%	4.16%
Sep 1983	4.85%	4.43%	4.27%
Oct 1983	4.60%	4.31%	3.90%
Nov 1983	3.69%	3.36%	3.36%
Dec 1983	4.21%	3.86%	3.54%
Jan 1984	4.92%	4.68%	4.18%
Feb 1984	5.19%	5.00%	4.36%
Mar 1984	5.72%	5.35%	4.45%
Apr 1984	5.48%	5.33%	4.23%
May 1984	5.36%	5.26%	4.30%
Jun 1984	4.39%	4.47%	3.40%
Average Premiums	4.83%	4.56%	4.01%

Exhibit 6. Utility Risk Premiums and Interest Rates, 1980–1984

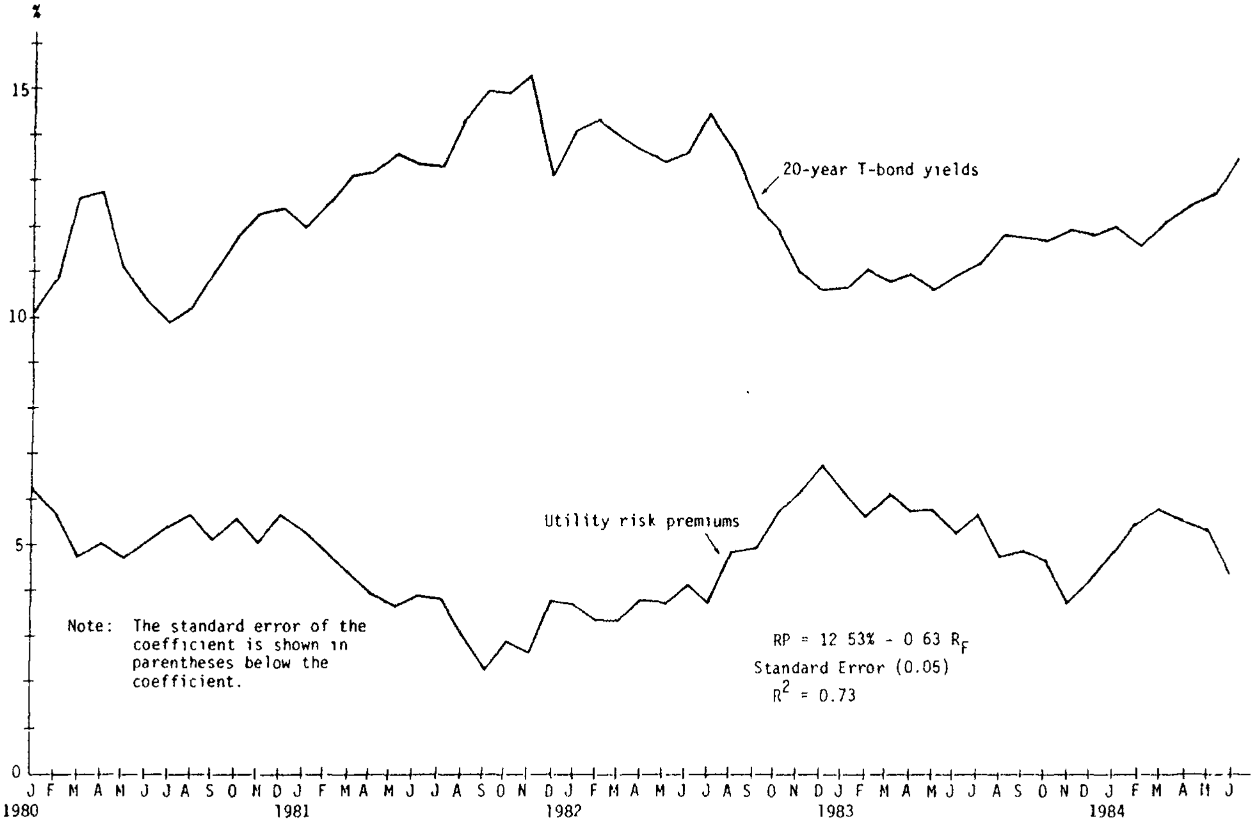


Exhibit 7. Monthly Risk Premiums, Electric Utilities, 1981–1984 (to Date)

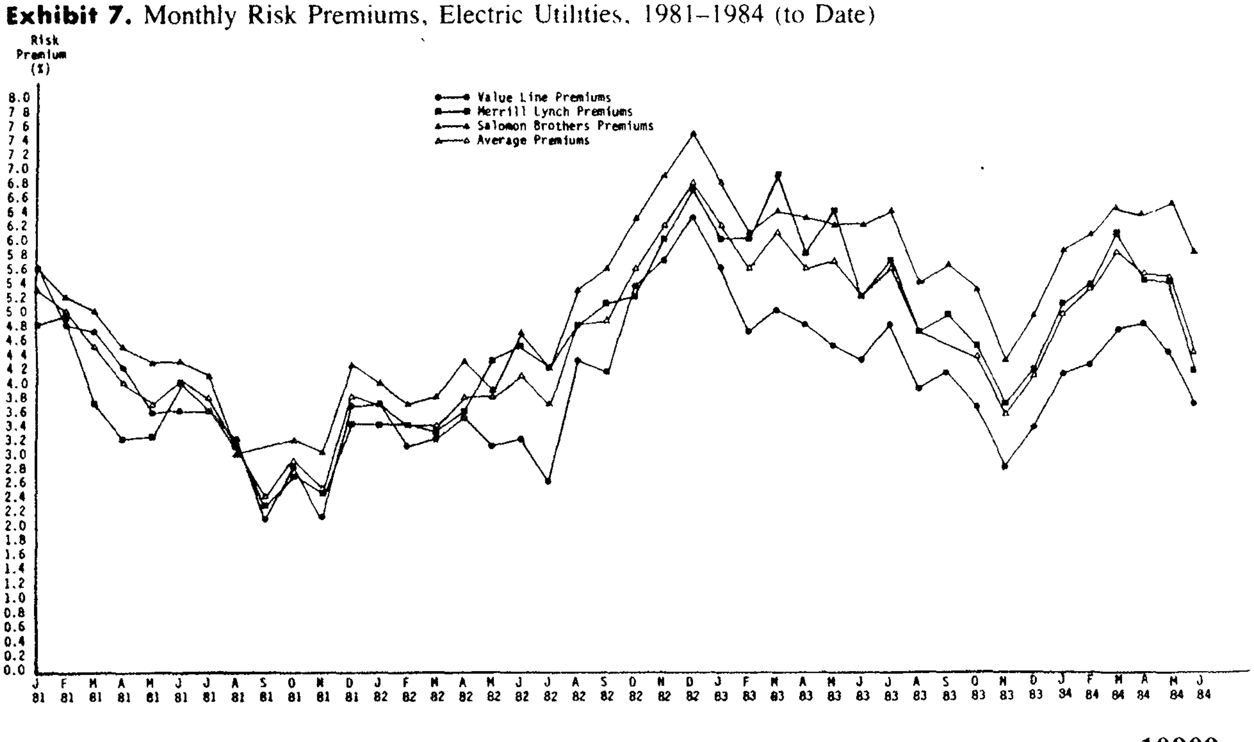
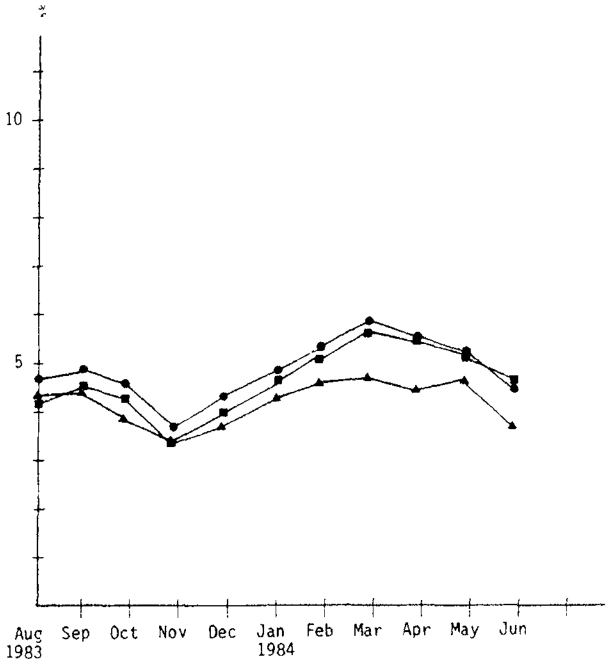


Exhibit 8. Comparative Risk Premium Data



●: Value Line, ML, SB: Dow Jones Electrics
■: IBES: Dow Jones Electrics
▲: IBES: All Electric Utilities

do differ, the differences are not large given the nature of the estimates, and the premiums follow one another closely over time. Since all of the analysts are examining essentially the same data and since utility companies are not competitive with one another, and hence have relatively few secrets, the similarity among the analysts' forecasts is not surprising.

4. The IBES data, presented in Exhibit 5 and plotted in Exhibit 8, contain too few observations to enable us to draw strong conclusions, but (i) the Dow Jones Electrics risk premiums based on our three-analyst data have averaged 27 basis points above premiums based on the larger group of analysts surveyed by IBES and (ii) the premiums on the 11 Dow Jones Electrics have averaged 54 basis points higher than premiums for the entire utility industry followed by IBES. Given the variability in the data, we are, at this point, inclined to attribute these differences to random fluctuations, but as more data become available, it may turn out that the differences are statistically significant. In particular, the 11 electric utilities included in the Dow Jones Utility Index all have large nuclear investments, and this may cause them to be regarded as riskier than the industry average, which includes both nuclear and non-nuclear companies.

Tests of the Reasonableness of the Risk Premium Estimates

So far our claims to the reasonableness of our risk-premium estimates have been based on the reasonableness of our variable measures, particularly the measures of expected dividend growth rates. Essentially, we have argued that since there is strong evidence in the literature in support of analysts' forecasts, risk premiums based on these forecasts are reasonable. In the spirit of positive economics, however, it is also important to demonstrate the reasonableness of our results more directly.

It is theoretically possible to test for the validity of the risk-premium estimates in a CAPM framework. In a cross-sectional estimate of the CAPM equation,

$$(k - R_f)_i = \alpha_0 + \alpha_1\beta_i + u_i, \qquad (5)$$

we would expect

$$\hat{\alpha}_0 = 0 \text{ and } \hat{\alpha}_1 = k_M - R_f = \text{Market risk premium.}$$

This test, of course, would be a joint test of both the CAPM and the reasonableness of our risk-premium estimates. There is a great deal of evidence that questions the empirical validity of the CAPM, especially when applied to regulated utilities. Under these conditions, it is obvious that no unambiguous conclusion can be drawn regarding the efficacy of the premium estimates from such a test.[8]

A simpler and less ambiguous test is to show that the risk premiums are higher for lower rated firms than for higher rated firms. Using 1984 data, we classified the

[8]We carried out the test on a monthly basis for 1984 and found positive but statistically insignificant coefficients. A typical result (for April 1984) follows:

$$(k - R_f)_i = \underset{(0.91)}{3.1675} + \underset{(1.44)}{1.8031}\ \beta_i$$

The figures in parentheses are standard errors. Utility risk premiums do increase with betas, but the intercept term is not zero as the CAPM would predict, and α_1 is both less than the predicted value and not statistically significant. Again, the observation that the coefficients do not conform to CAPM predictions could be as much a problem with CAPM specification for utilities as with the risk premium estimates.

A similar test was carried out by Friend, Westerfield, and Granito [9]. They tested the CAPM using expectational (survey) data rather than *ex post* holding period returns. They actually found their coefficient of β_i to be negative in all their cross-sectional tests.

Exhibit 9. Relationship between Risk Premiums and Bond Ratings, 1984*

Month	Aaa/AA	AA	Aa/A	A	A/BBB	BBB	Below BBB
January†	—	2.61%	3.06%	3.70%	5.07%	4.90%	9.45%
February	2.98%	3.17%	3.36%	4.03%	5.26%	5.14%	7.97%
March	2.34%	3.46%	3.29%	4.06%	5.43%	5.02%	8.28%
April	2.37%	3.03%	3.29%	3.88%	5.29%	4.97%	6.96%
May	2.00%	2.48%	3.42%	3.72%	4.72%	6.64%	8.81%
June	0.72%	2.17%	2.46%	3.16%	3.76%	5.00%	5.58%
Average	2.08%	2.82%	3.15%	3.76%	4.92%	5.28%	7.84%

*The risk premiums are based on IBES data for the electric utilities followed by both IBES and Salomon Brothers. The number of electric utilities followed by both firms varies from month to month. For the period between January and June 1984, the number of electrics followed by both firms ranged from 96 to 99 utilities.

†In January, there were no Aaa/AA companies. Subsequently, four utilities were upgraded to Aaa/AA.

utility industry into risk groups based on bond ratings. For each rating group, we estimated the average risk premium. The results, presented in Exhibit 9, clearly show that the lower the bond rating, the higher the risk premiums. Our premium estimates therefore would appear to pass this simple test of reasonableness.

Risk Premiums and Interest Rates

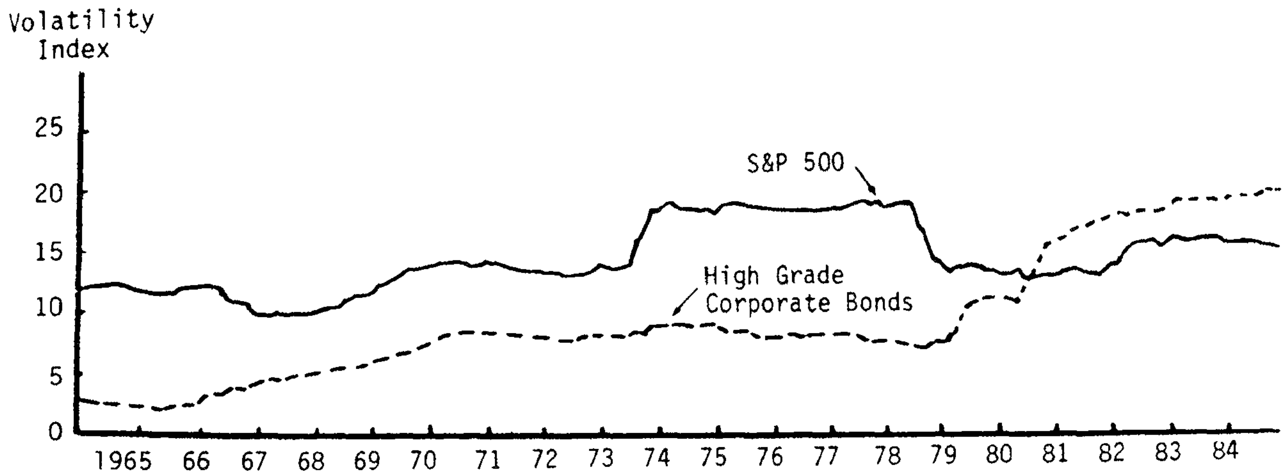
Traditionally, stocks have been regarded as being riskier than bonds because bondholders have a prior claim on earnings and assets. That is, stockholders stand at the end of the line and receive income and/or assets only after the claims of bondholders have been satisfied. However, if interest rates fluctuate, then the holders of long-term bonds can suffer losses (either realized or in an opportunity cost sense) even though they receive all contractually due payments. Therefore, if investors' worries about "interest rate risk" versus "earning power risk" vary over time, then perceived risk differentials between stocks and bonds, and hence risk premiums, will also vary.

Any number of events could occur to cause the perceived riskiness of stocks versus bonds to change, but probably the most pervasive factor, over the 1966–1984 period, is related to inflation. Inflationary expectations are, of course, reflected in interest rates. Therefore, one might expect to find a relationship between risk premiums and interest rates. As we noted in our discussion of Exhibit 3, risk premiums were positively correlated with interest rates from 1966 through 1979, but, beginning in 1980, the relationship turned negative. A possible explanation for this change is given next.

1966–1979 Period. During this period, inflation heated up, fuel prices soared, environmental problems surfaced, and demand for electricity slowed even as expensive new generating units were nearing completion. These cost increases required offsetting rate hikes to maintain profit levels. However, political pressure, combined with administrative procedures that were not designed to deal with a volatile economic environment, led to long periods of "regulatory lag" that caused utilities' earned ROEs to decline in absolute terms and to fall far below the cost of equity. These factors combined to cause utility stockholders to experience huge losses: S&P's Electric Index dropped from a mid-1960s high of 60.90 to a mid-1970s low of 20.41, a decrease of 66.5%. Industrial stocks also suffered losses during this period, but, on average, they were only one third as severe as the utilities' losses. Similarly, investors in long-term bonds had losses, but bond losses were less than half those of utility stocks. Note also that, during this period, (i) bond investors were able to reinvest coupons and maturity payments at rising rates, whereas the earned returns on equity did not rise, and (ii) utilities were providing a rising share of their operating income to debtholders versus stockholders (interest expense/book value of debt was rising, while net income/common equity was declining). This led to a widespread belief that utility commissions would provide enough revenues to keep utilities from going bankrupt (barring a disaster), and hence to protect the bondholders, but that they would not necessarily provide enough revenues either to permit the expected rate of dividend growth to occur or, perhaps, even to allow the dividend to be maintained.

Because of these experiences, investors came to regard inflation as having a more negative effect on utility stocks than on bonds. Therefore, when fears of inflation increased, utilities' measured risk premiums

Exhibit 10. Relative Volatility* of Stocks and Bonds, 1965–1984



*Volatility is measured as the standard deviation of total returns over the last 5 years
Source Merrill Lynch, *Quantitative Analysis*, May/June 1984

also increased. A regression over the period 1966–1979, using our Exhibit 2 data, produced this result.

$$RP = 0.30\% + \underset{(0.22)}{0.73}\ R_{T};\qquad r^2 = 0.48.$$

This indicates that a one percentage point increase in the Treasury bond rate produced, on average, a 0.73 percentage point increase in the risk premium, and hence a 1.00 + 0.73 = 1.73 percentage point increase in the cost of equity for utilities.

1980–1984 Period. The situation changed dramatically in 1980 and thereafter. Except for a few companies with nuclear construction problems, the utilities' financial situations stabilized in the early 1980s, and then improved significantly from 1982 to 1984. Both the companies and their regulators were learning to live with inflation; many construction programs were completed; regulatory lags were shortened; and in general the situation was much better for utility equity investors. In the meantime, over most of the 1980–1984 period, interest rates and bond prices fluctuated violently, both in an absolute sense and relative to common stocks. Exhibit 10 shows the volatility of corporate bonds very clearly. Over most of the eighteen-year period, stock returns were much more volatile than returns on bonds. However, that situation changed in October 1979, when the Fed began to focus on the money supply rather than on interest rates.[9]

In the 1980–1984 period, an increase in inflationary expectations has had a more adverse effect on bonds than on utility stocks. If the expected rate of inflation increases, then interest rates *will increase* and bond prices *will fall* Thus, uncertainty about inflation translates directly into risk in the bond markets The effect of inflation on stocks, including utility stocks, is less clear. If inflation increases, then utilities should, in theory, be able to obtain rate increases that would offset increases in operating costs and also compensate for the higher cost of equity. Thus, with "proper" regulation, utility stocks would provide a better hedge against unanticipated inflation than would bonds. This hedge did not work at all well during the 1966–1979 period, because inflation-induced increases in operating and capital costs were not offset by timely rate increases. However, as noted earlier, both the utilities and their regulators seem to have learned to live better with inflation during the 1980s.

Since inflation is today regarded as a major investment risk, and since utility stocks now seem to provide a better hedge against unanticipated inflation than do

[9]Because the standard deviations in Exhibit 10 are based on the last five years of data, even if bond returns stabilize, as they did beginning in 1982, their reported volatility will remain high for several more years Thus, Exhibit 10 gives a rough indication of the current relative riskiness of stocks versus bonds, but the measure is by no means precise or necessarily indicative of future expectations

bonds, the interest-rate risk inherent in bonds offsets, to a greater extent than was true earlier, the higher operating risk that is inherent in equities. Therefore, when inflationary fears rise, the perceived riskiness of bonds rises, helping to push up interest rates. However, since investors are today less concerned about inflation's impact on utility stocks than on bonds, the utilities' cost of equity does not rise as much as that of debt, so the observed risk premium tends to fall.

For the 1980–1984 period, we found the following relationship (see Exhibit 6):

$$RP = 12.53\% - \underset{(0.05)}{0.63} R_F; \qquad r^2 = 0.73.$$

Thus, a one percentage point increase in the T-bond rate, on average, caused the risk premium to fall by 0.63%, and hence it led to a 1.00 − 0.63 = 0.37 percentage point increase in the cost of equity to an average utility. This contrasts sharply with the pre-1980 period, when a one percentage point increase in interest rates led, on average, to a 1.73 percentage point increase in the cost of equity.

Summary and Implications

We began by reviewing a number of earlier studies. From them, we concluded that, for cost of capital estimation purposes, risk premiums must be based on expectations, not on past realized holding period returns. Next, we noted that expectational risk premiums may be estimated either from surveys, such as the ones Charles Benore has conducted, or by use of DCF techniques. Further, we found that, although growth rates for use in the DCF model can be either developed from time-series data or obtained from security analysts, analysts' growth forecasts are more reflective of investors' views, and, hence, in our opinion are preferable for use in risk-premium studies.

Using analysts' growth rates and the DCF model, we estimated risk premiums over several different periods. From 1966 to 1984, risk premiums for both electric utilities and industrial stocks varied widely from year to year. Also, during the first half of the period, the utilities had smaller risk premiums than the industrials, but after the mid-1970s, the risk premiums for the two groups were, on average, about equal.

The effects of changing interest rates on risk premiums shifted dramatically in 1980, at least for the utilities. From 1965 through 1979, inflation generally had a more severe adverse effect on utility stocks than on bonds, and, as a result, an increase in inflationary expectations, as reflected in interest rates, caused an increase in equity risk premiums. However, in 1980 and thereafter, rising inflation and interest rates increased the perceived riskiness of bonds more than that of utility equities, so the relationship between interest rates and utility risk premiums shifted from positive to negative. Earlier, a 1.00 percentage point increase in interest rates had led, on average, to a 1.73% increase in the utilities' cost of equity, but after 1980 a 1.00 percentage point increase in the cost of debt was associated with an increase of only 0.37% in the cost of equity.

Our study also has implications for the use of the CAPM to estimate the cost of equity for utilities. The CAPM studies that we have seen typically use either Ibbotson-Sinquefield or similar historic holding period returns as the basis for estimating the market risk premium. Such usage implicitly assumes (i) that *ex post* returns data can be used to proxy *ex ante* expectations and (ii) that the market risk premium is relatively stable over time. Our analysis suggests that neither of these assumptions is correct; at least for utility stocks, *ex post* returns data do not appear to be reflective of *ex ante* expectations, and risk premiums are volatile, not stable.

Unstable risk premiums also make us question the FERC and FCC proposals to estimate a risk premium for the utilities every two years and then to add this premium to a current Treasury bond rate to determine a utility's cost of equity. Administratively, this proposal would be easy to handle, but risk premiums are simply too volatile to be left in place for two years.

References

1. C. Benore, *A Survey of Investor Attitudes toward the Electric Power Industry*, New York, Paine Webber Mitchell Hutchins, Inc., May 1983.
2. E. F. Brigham and D. K. Shome, "The Risk Premium Approach to Estimating the Cost of Common Equity Capital," *Proceedings of the Iowa State Regulatory Conference* (May 1980), pp. 239–275.
3. ——— "Estimating the Market Risk Premium," in R. L. Crum and F. G. J. Derkinderin (eds.), *Risk, Capital Costs, and Project Financing Decisions*, Nijenrode Studies in Business, Boston, Martinus Nijhoff, 1981.
4. ———"Equity Risk Premiums in the 1980s," in *Earnings Regulation under Inflation*, Washington, DC, Institute for the Study of Regulation, 1982, pp. 166–181.
5. L. D. Brown and M.S. Rozeff, "The Superiority of Analysts' Forecasts as a Measure of Expectations: Evidence from Earnings," *Journal of Finance* (March 1978), pp. 1–16.

6. W T. Carleton, D R. Chambers, and J Lakonishok, "Inflation Risk and Regulatory Lag," *Journal of Finance* (May 1983), pp. 419–431.
7 J.G. Cragg and B.G Malkiel, *Expectations and the Structure of Share Prices*, Chicago, The University of Chicago Press, 1982
8 E F. Fama and W.G. Schwert, "Asset Returns and Inflation," *Journal of Financial Economics*, November 1977, pp. 115–146.
9. I Friend, R. Westerfield, and M. Granito, "New Evidence on the Capital Asset Pricing Model," *Journal of Finance* (June 1978), pp 903–917.
10. M.J Gordon and P J. Halpern, "Bond Share Yield Spreads under Uncertain Inflation," *American Economic Review* (September 1976), pp. 559–565.
11 N B Gultekin, "Stock Market Returns and Inflation Forecasts," *Journal of Finance* (June 1983), pp 663–673.
12 R.G. Ibbotson and R.A. Sinquefield, *Stocks, Bonds, Bills, and Inflation: Historical Returns (1926–1978)*, Charlottesville, VA, Financial Analysts Research Foundation, 1979
13. C M. Linke, "Estimating Growth Expectations for AT&T. Survey Approach," Washington, DC, Advanced Seminar on Earnings Regulation, November 1981.
14. B G Malkiel, "The Capital Formation Problem in the United States," *Journal of Finance*, May 1979, pp 291–306
15. A A Robichek, "Regulation and Modern Finance Theory," *Journal of Finance* (June 1978), pp 693–705
16 K L Stanley, W G Lewellen, and G G. Schlarbaum, "Further Evidence on the Value of Professional Investment Research," *Journal of Financial Research* (Spring 1981), pp 1–9
17 Touche, Ross, and Company, *Proxy Disclosures and Stockholder Attitude Survey*, Washington, DC, National Association of Corporate Directors, May 1982
18 R.F Vandell and G.W Kester, *A History of Risk Premia Estimates for Equities 1944–1978*, Charlottesville, VA, Financial Analysts Research Foundation, 1983

Journal of Economic Perspectives—Volume 18, Number 3—Summer 2004—Pages 25–46

The Capital Asset Pricing Model: Theory and Evidence

Eugene F. Fama and Kenneth R. French

The capital asset pricing model (CAPM) of William Sharpe (1964) and John Lintner (1965) marks the birth of asset pricing theory (resulting in a Nobel Prize for Sharpe in 1990). Four decades later, the CAPM is still widely used in applications, such as estimating the cost of capital for firms and evaluating the performance of managed portfolios. It is the centerpiece of MBA investment courses. Indeed, it is often the only asset pricing model taught in these courses.[1]

The attraction of the CAPM is that it offers powerful and intuitively pleasing predictions about how to measure risk and the relation between expected return and risk. Unfortunately, the empirical record of the model is poor—poor enough to invalidate the way it is used in applications. The CAPM's empirical problems may reflect theoretical failings, the result of many simplifying assumptions. But they may also be caused by difficulties in implementing valid tests of the model. For example, the CAPM says that the risk of a stock should be measured relative to a comprehensive "market portfolio" that in principle can include not just traded financial assets, but also consumer durables, real estate and human capital. Even if we take a narrow view of the model and limit its purview to traded financial assets, is it

[1] Although every asset pricing model is a capital asset pricing model, the finance profession reserves the acronym CAPM for the specific model of Sharpe (1964), Lintner (1965) and Black (1972) discussed here. Thus, throughout the paper we refer to the Sharpe-Lintner-Black model as the CAPM.

■ *Eugene F. Fama is Robert R. McCormick Distinguished Service Professor of Finance, Graduate School of Business, University of Chicago, Chicago, Illinois. Kenneth R. French is Carl E. and Catherine M. Heidt Professor of Finance, Tuck School of Business, Dartmouth College, Hanover, New Hampshire. Their e-mail addresses are ⟨eugene.fama@gsb.uchicago.edu⟩ and ⟨kfrench@dartmouth.edu⟩, respectively.*

legitimate to limit further the market portfolio to U.S. common stocks (a typical choice), or should the market be expanded to include bonds, and other financial assets, perhaps around the world? In the end, we argue that whether the model's problems reflect weaknesses in the theory or in its empirical implementation, the failure of the CAPM in empirical tests implies that most applications of the model are invalid.

We begin by outlining the logic of the CAPM, focusing on its predictions about risk and expected return. We then review the history of empirical work and what it says about shortcomings of the CAPM that pose challenges to be explained by alternative models.

The Logic of the CAPM

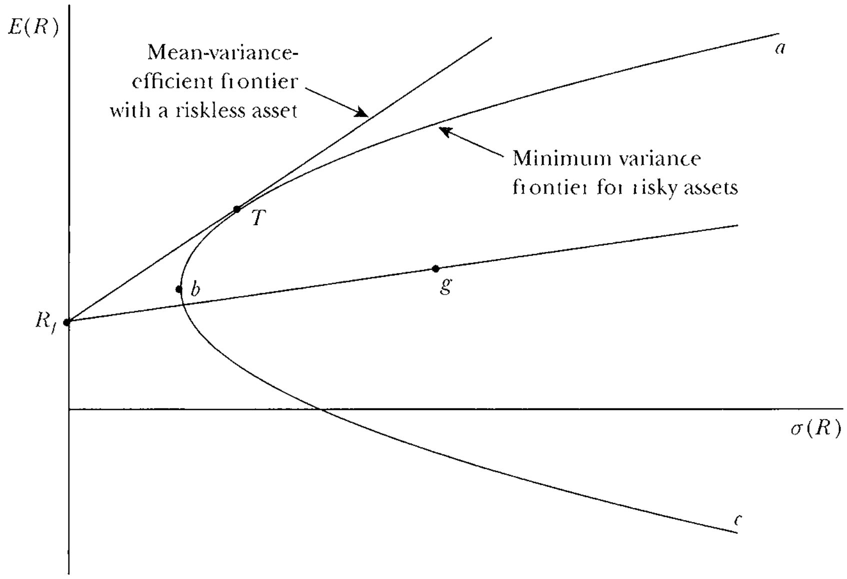
The CAPM builds on the model of portfolio choice developed by Harry Markowitz (1959). In Markowitz's model, an investor selects a portfolio at time $t - 1$ that produces a stochastic return at t. The model assumes investors are risk averse and, when choosing among portfolios, they care only about the mean and variance of their one-period investment return. As a result, investors choose "mean-variance-efficient" portfolios, in the sense that the portfolios 1) minimize the variance of portfolio return, given expected return, and 2) maximize expected return, given variance. Thus, the Markowitz approach is often called a "mean-variance model."

The portfolio model provides an algebraic condition on asset weights in mean-variance-efficient portfolios. The CAPM turns this algebraic statement into a testable prediction about the relation between risk and expected return by identifying a portfolio that must be efficient if asset prices are to clear the market of all assets.

Sharpe (1964) and Lintner (1965) add two key assumptions to the Markowitz model to identify a portfolio that must be mean-variance-efficient. The first assumption is *complete agreement*: given market clearing asset prices at $t - 1$, investors agree on the joint distribution of asset returns from $t - 1$ to t. And this distribution is the true one—that is, it is the distribution from which the returns we use to test the model are drawn. The second assumption is that there is *borrowing and lending at a risk-free rate*, which is the same for all investors and does not depend on the amount borrowed or lent.

Figure 1 describes portfolio opportunities and tells the CAPM story. The horizontal axis shows portfolio risk, measured by the standard deviation of portfolio return; the vertical axis shows expected return. The curve *abc*, which is called the minimum variance frontier, traces combinations of expected return and risk for portfolios of risky assets that minimize return variance at different levels of expected return. (These portfolios do not include risk-free borrowing and lending.) The tradeoff between risk and expected return for minimum variance portfolios is apparent. For example, an investor who wants a high expected return, perhaps at point a, must accept high volatility. At point T, the investor can have an interme-

Figure 1
Investment Opportunities



diate expected return with lower volatility. If there is no risk-free borrowing or lending, only portfolios above b along abc are mean-variance-efficient, since these portfolios also maximize expected return, given their return variances.

Adding risk-free borrowing and lending turns the efficient set into a straight line. Consider a portfolio that invests the proportion x of portfolio funds in a risk-free security and $1 - x$ in some portfolio g. If all funds are invested in the risk-free security—that is, they are loaned at the risk-free rate of interest—the result is the point R_f in Figure 1, a portfolio with zero variance and a risk-free rate of return. Combinations of risk-free lending and positive investment in g plot on the straight line between R_f and g. Points to the right of g on the line represent borrowing at the risk-free rate, with the proceeds from the borrowing used to increase investment in portfolio g. In short, portfolios that combine risk-free lending or borrowing with some risky portfolio g plot along a straight line from R_f through g in Figure 1.[2]

[2] Formally, the return, expected return and standard deviation of return on portfolios of the risk-free asset f and a risky portfolio g vary with x, the proportion of portfolio funds invested in f, as

$$R_p = xR_f + (1 - x)R_g,$$

$$E(R_p) = xR_f + (1 - x)E(R_g),$$

$$\sigma(R_p) = (1 - x)\sigma(R_g),\ x \leq 1.0,$$

which together imply that the portfolios plot along the line from R_f through g in Figure 1.

To obtain the mean-variance-efficient portfolios available with risk-free borrowing and lending, one swings a line from R_f in Figure 1 up and to the left as far as possible, to the tangency portfolio T. We can then see that all efficient portfolios are combinations of the risk-free asset (either risk-free borrowing or lending) and a single risky tangency portfolio, T. This key result is Tobin's (1958) "separation theorem."

The punch line of the CAPM is now straightforward. With complete agreement about distributions of returns, all investors see the same opportunity set (Figure 1), and they combine the same risky tangency portfolio T with risk-free lending or borrowing. Since all investors hold the same portfolio T of risky assets, it must be the value-weight market portfolio of risky assets. Specifically, each risky asset's weight in the tangency portfolio, which we now call M (for the "market"), must be the total market value of all outstanding units of the asset divided by the total market value of all risky assets. In addition, the risk-free rate must be set (along with the prices of risky assets) to clear the market for risk-free borrowing and lending.

In short, the CAPM assumptions imply that the market portfolio M must be on the minimum variance frontier if the asset market is to clear. This means that the algebraic relation that holds for any minimum variance portfolio must hold for the market portfolio. Specifically, if there are N risky assets,

$$\text{(Minimum Variance Condition for } M\text{)} \quad E(R_i) = E(R_{ZM}) + [E(R_M) - E(R_{ZM})]\beta_{iM},\ i = 1, \ldots, N.$$

In this equation, $E(R_i)$ is the expected return on asset i, and β_{iM}, the market beta of asset i, is the covariance of its return with the market return divided by the variance of the market return,

$$\text{(Market Beta)} \quad \beta_{iM} = \frac{\text{cov}(R_i, R_M)}{\sigma^2(R_M)}.$$

The first term on the right-hand side of the minimum variance condition, $E(R_{ZM})$, is the expected return on assets that have market betas equal to zero, which means their returns are uncorrelated with the market return. The second term is a risk premium—the market beta of asset i, β_{iM}, times the premium per unit of beta, which is the expected market return, $E(R_M)$, minus $E(R_{ZM})$.

Since the market beta of asset i is also the slope in the regression of its return on the market return, a common (and correct) interpretation of beta is that it measures the sensitivity of the asset's return to variation in the market return. But there is another interpretation of beta more in line with the spirit of the portfolio model that underlies the CAPM. The risk of the market portfolio, as measured by the variance of its return (the denominator of β_{iM}), is a weighted average of the covariance risks of the assets in M (the numerators of β_{iM} for different assets).

Thus, β_{iM} is the covariance risk of asset i in M measured relative to the average covariance risk of assets, which is just the variance of the market return.[3] In economic terms, β_{iM} is proportional to the risk each dollar invested in asset i contributes to the market portfolio.

The last step in the development of the Sharpe-Lintner model is to use the assumption of risk-free borrowing and lending to nail down $E(R_{ZM})$, the expected return on zero-beta assets. A risky asset's return is uncorrelated with the market return—its beta is zero—when the average of the asset's covariances with the returns on other assets just offsets the variance of the asset's return. Such a risky asset is riskless in the market portfolio in the sense that it contributes nothing to the variance of the market return.

When there is risk-free borrowing and lending, the expected return on assets that are uncorrelated with the market return, $E(R_{ZM})$, must equal the risk-free rate, R_f. The relation between expected return and beta then becomes the familiar Sharpe-Lintner CAPM equation,

$$\text{(Sharpe-Lintner CAPM)} \quad E(R_i) = R_f + [E(R_M) - R_f)]\beta_{iM}, \; i = 1, \ldots, N.$$

In words, the expected return on any asset i is the risk-free interest rate, R_f, plus a risk premium, which is the asset's market beta, β_{iM}, times the premium per unit of beta risk, $E(R_M) - R_f$.

Unrestricted risk-free borrowing and lending is an unrealistic assumption. Fischer Black (1972) develops a version of the CAPM without risk-free borrowing or lending. He shows that the CAPM's key result—that the market portfolio is mean-variance-efficient—can be obtained by instead allowing unrestricted short sales of risky assets. In brief, back in Figure 1, if there is no risk-free asset, investors select portfolios from along the mean-variance-efficient frontier from *a* to *b*. Market clearing prices imply that when one weights the efficient portfolios chosen by investors by their (positive) shares of aggregate invested wealth, the resulting portfolio is the market portfolio. The market portfolio is thus a portfolio of the efficient portfolios chosen by investors. With unrestricted short selling of risky assets, portfolios made up of efficient portfolios are themselves efficient. Thus, the market portfolio is efficient, which means that the minimum variance condition for M given above holds, and it is the expected return-risk relation of the Black CAPM.

The relations between expected return and market beta of the Black and Sharpe-Lintner versions of the CAPM differ only in terms of what each says about $E(R_{ZM})$, the expected return on assets uncorrelated with the market. The Black version says only that $E(R_{ZM})$ must be less than the expected market return, so the

[3] Formally, if x_{iM} is the weight of asset i in the market portfolio, then the variance of the portfolio's return is

$$\sigma^2(R_M) = Cov(R_M, R_M) = Cov\left(\sum_{i=1}^{N} x_{iM}R_i,\ R_M\right) = \sum_{i=1}^{N} x_{iM}Cov(R_i, R_M).$$

premium for beta is positive. In contrast, in the Sharpe-Lintner version of the model, $E(R_{ZM})$ must be the risk-free interest rate, R_f, and the premium per unit of beta risk is $E(R_M) - R_f$.

The assumption that short selling is unrestricted is as unrealistic as unrestricted risk-free borrowing and lending. If there is no risk-free asset and short sales of risky assets are not allowed, mean-variance investors still choose efficient portfolios—points above *b* on the *abc* curve in Figure 1. But when there is no short selling of risky assets and no risk-free asset, the algebra of portfolio efficiency says that portfolios made up of efficient portfolios are not typically efficient. This means that the market portfolio, which is a portfolio of the efficient portfolios chosen by investors, is not typically efficient. And the CAPM relation between expected return and market beta is lost. This does not rule out predictions about expected return and betas with respect to other efficient portfolios—if theory can specify portfolios that must be efficient if the market is to clear. But so far this has proven impossible.

In short, the familiar CAPM equation relating expected asset returns to their market betas is just an application to the market portfolio of the relation between expected return and portfolio beta that holds in any mean-variance-efficient portfolio. The efficiency of the market portfolio is based on many unrealistic assumptions, including complete agreement and either unrestricted risk-free borrowing and lending or unrestricted short selling of risky assets. But all interesting models involve unrealistic simplifications, which is why they must be tested against data.

Early Empirical Tests

Tests of the CAPM are based on three implications of the relation between expected return and market beta implied by the model. First, expected returns on all assets are linearly related to their betas, and no other variable has marginal explanatory power. Second, the beta premium is positive, meaning that the expected return on the market portfolio exceeds the expected return on assets whose returns are uncorrelated with the market return. Third, in the Sharpe-Lintner version of the model, assets uncorrelated with the market have expected returns equal to the risk-free interest rate, and the beta premium is the expected market return minus the risk-free rate. Most tests of these predictions use either cross-section or time-series regressions. Both approaches date to early tests of the model.

Tests on Risk Premiums

The early cross-section regression tests focus on the Sharpe-Lintner model's predictions about the intercept and slope in the relation between expected return and market beta. The approach is to regress a cross-section of average asset returns on estimates of asset betas. The model predicts that the intercept in these regressions is the risk-free interest rate, R_f, and the coefficient on beta is the expected return on the market in excess of the risk-free rate, $E(R_M) - R_f$.

Two problems in these tests quickly became apparent. First, estimates of beta for individual assets are imprecise, creating a measurement error problem when they are used to explain average returns. Second, the regression residuals have common sources of variation, such as industry effects in average returns. Positive correlation in the residuals produces downward bias in the usual ordinary least squares estimates of the standard errors of the cross-section regression slopes.

To improve the precision of estimated betas, researchers such as Blume (1970), Friend and Blume (1970) and Black, Jensen and Scholes (1972) work with portfolios, rather than individual securities. Since expected returns and market betas combine in the same way in portfolios, if the CAPM explains security returns it also explains portfolio returns.[4] Estimates of beta for diversified portfolios are more precise than estimates for individual securities. Thus, using portfolios in cross-section regressions of average returns on betas reduces the critical errors in variables problem. Grouping, however, shrinks the range of betas and reduces statistical power. To mitigate this problem, researchers sort securities on beta when forming portfolios; the first portfolio contains securities with the lowest betas, and so on, up to the last portfolio with the highest beta assets. This sorting procedure is now standard in empirical tests.

Fama and MacBeth (1973) propose a method for addressing the inference problem caused by correlation of the residuals in cross-section regressions. Instead of estimating a single cross-section regression of average monthly returns on betas, they estimate month-by-month cross-section regressions of monthly returns on betas. The times-series means of the monthly slopes and intercepts, along with the standard errors of the means, are then used to test whether the average premium for beta is positive and whether the average return on assets uncorrelated with the market is equal to the average risk-free interest rate. In this approach, the standard errors of the average intercept and slope are determined by the month-to-month variation in the regression coefficients, which fully captures the effects of residual correlation on variation in the regression coefficients, but sidesteps the problem of actually estimating the correlations. The residual correlations are, in effect, captured via repeated sampling of the regression coefficients. This approach also becomes standard in the literature.

Jensen (1968) was the first to note that the Sharpe-Lintner version of the

[4] Formally, if x_{ip}, $i = 1, \ldots, N$, are the weights for assets in some portfolio p, the expected return and market beta for the portfolio are related to the expected returns and betas of assets as

$$E(R_p) = \sum_{i=1}^{N} x_{ip}E(R_i), \text{ and } \beta_{pM} = \sum_{i=1}^{N} x_{ip}\beta_{pM}.$$

Thus, the CAPM relation between expected return and beta,

$$E(R_i) = E(R_f) + [E(R_M) - E(R_f)]\beta_{iM},$$

holds when asset i is a portfolio, as well as when i is an individual security.

relation between expected return and market beta also implies a time-series regression test. The Sharpe-Lintner CAPM says that the expected value of an asset's excess return (the asset's return minus the risk-free interest rate, $R_{it} - R_{ft}$) is completely explained by its expected CAPM risk premium (its beta times the expected value of $R_{Mt} - R_{ft}$). This implies that "Jensen's alpha," the intercept term in the time-series regression,

$$\text{(Time-Series Regression)} \quad R_{it} - R_{ft} = \alpha_i + \beta_{iM}(R_{Mt} - R_{ft}) + \varepsilon_{it},$$

is zero for each asset.

The early tests firmly reject the Sharpe-Lintner version of the CAPM. There is a positive relation between beta and average return, but it is too "flat." Recall that, in cross-section regressions, the Sharpe-Lintner model predicts that the intercept is the risk-free rate and the coefficient on beta is the expected market return in excess of the risk-free rate, $E(R_M) - R_f$. The regressions consistently find that the intercept is greater than the average risk-free rate (typically proxied as the return on a one-month Treasury bill), and the coefficient on beta is less than the average excess market return (proxied as the average return on a portfolio of U.S. common stocks minus the Treasury bill rate). This is true in the early tests, such as Douglas (1968), Black, Jensen and Scholes (1972), Miller and Scholes (1972), Blume and Friend (1973) and Fama and MacBeth (1973), as well as in more recent cross-section regression tests, like Fama and French (1992).

The evidence that the relation between beta and average return is too flat is confirmed in time-series tests, such as Friend and Blume (1970), Black, Jensen and Scholes (1972) and Stambaugh (1982). The intercepts in time-series regressions of excess asset returns on the excess market return are positive for assets with low betas and negative for assets with high betas.

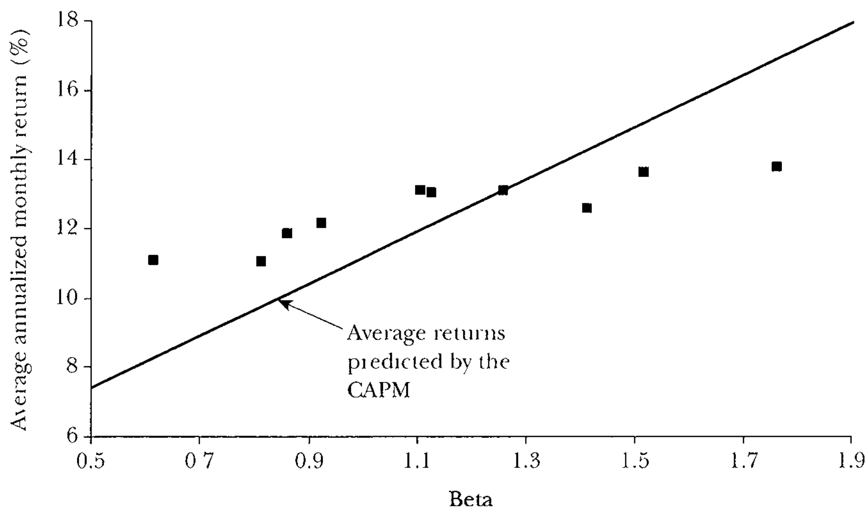
Figure 2 provides an updated example of the evidence. In December of each year, we estimate a preranking beta for every NYSE (1928–2003), AMEX (1963–2003) and NASDAQ (1972–2003) stock in the CRSP (Center for Research in Security Prices of the University of Chicago) database, using two to five years (as available) of prior monthly returns.[5] We then form ten value-weight portfolios based on these preranking betas and compute their returns for the next twelve months. We repeat this process for each year from 1928 to 2003. The result is 912 monthly returns on ten beta-sorted portfolios. Figure 2 plots each portfolio's average return against its postranking beta, estimated by regressing its monthly returns for 1928–2003 on the return on the CRSP value-weight portfolio of U.S. common stocks.

The Sharpe-Lintner CAPM predicts that the portfolios plot along a straight

[5] To be included in the sample for year t, a security must have market equity data (price times shares outstanding) for December of $t - 1$, and CRSP must classify it as ordinary common equity. Thus, we exclude securities such as American Depository Receipts (ADRs) and Real Estate Investment Trusts (REITs).

Figure 2

Average Annualized Monthly Return versus Beta for Value Weight Portfolios Formed on Prior Beta, 1928–2003



line, with an intercept equal to the risk-free rate, R_f, and a slope equal to the expected excess return on the market, $E(R_M) - R_f$. We use the average one-month Treasury bill rate and the average excess CRSP market return for 1928–2003 to estimate the predicted line in Figure 2. Confirming earlier evidence, the relation between beta and average return for the ten portfolios is much flatter than the Sharpe-Lintner CAPM predicts. The returns on the low beta portfolios are too high, and the returns on the high beta portfolios are too low. For example, the predicted return on the portfolio with the lowest beta is 8.3 percent per year; the actual return is 11.1 percent. The predicted return on the portfolio with the highest beta is 16.8 percent per year; the actual is 13.7 percent.

Although the observed premium per unit of beta is lower than the Sharpe-Lintner model predicts, the relation between average return and beta in Figure 2 is roughly linear. This is consistent with the Black version of the CAPM, which predicts only that the beta premium is positive. Even this less restrictive model, however, eventually succumbs to the data.

Testing Whether Market Betas Explain Expected Returns

The Sharpe-Lintner and Black versions of the CAPM share the prediction that the market portfolio is mean-variance-efficient. This implies that differences in expected return across securities and portfolios are entirely explained by differences in market beta; other variables should add nothing to the explanation of expected return. This prediction plays a prominent role in tests of the CAPM. In the early work, the weapon of choice is cross-section regressions.

In the framework of Fama and MacBeth (1973), one simply adds predetermined explanatory variables to the month-by-month cross-section regressions of returns on beta. If all differences in expected return are explained by beta, the average slopes on the additional variables should not be reliably different from zero. Clearly, the trick in the cross-section regression approach is to choose specific additional variables likely to expose any problems of the CAPM prediction that, because the market portfolio is efficient, market betas suffice to explain expected asset returns.

For example, in Fama and MacBeth (1973) the additional variables are squared market betas (to test the prediction that the relation between expected return and beta is linear) and residual variances from regressions of returns on the market return (to test the prediction that market beta is the only measure of risk needed to explain expected returns). These variables do not add to the explanation of average returns provided by beta. Thus, the results of Fama and MacBeth (1973) are consistent with the hypothesis that their market proxy—an equal-weight portfolio of NYSE stocks—is on the minimum variance frontier.

The hypothesis that market betas completely explain expected returns can also be tested using time-series regressions. In the time-series regression described above (the excess return on asset i regressed on the excess market return), the intercept is the difference between the asset's average excess return and the excess return predicted by the Sharpe-Lintner model, that is, beta times the average excess market return. If the model holds, there is no way to group assets into portfolios whose intercepts are reliably different from zero. For example, the intercepts for a portfolio of stocks with high ratios of earnings to price and a portfolio of stocks with low earning-price ratios should both be zero. Thus, to test the hypothesis that market betas suffice to explain expected returns, one estimates the time-series regression for a set of assets (or portfolios) and then jointly tests the vector of regression intercepts against zero. The trick in this approach is to choose the left-hand-side assets (or portfolios) in a way likely to expose any shortcoming of the CAPM prediction that market betas suffice to explain expected asset returns.

In early applications, researchers use a variety of tests to determine whether the intercepts in a set of time-series regressions are all zero. The tests have the same asymptotic properties, but there is controversy about which has the best small sample properties. Gibbons, Ross and Shanken (1989) settle the debate by providing an F-test on the intercepts that has exact small-sample properties. They also show that the test has a simple economic interpretation. In effect, the test constructs a candidate for the tangency portfolio T in Figure 1 by optimally combining the market proxy and the left-hand-side assets of the time-series regressions. The estimator then tests whether the efficient set provided by the combination of this tangency portfolio and the risk-free asset is reliably superior to the one obtained by combining the risk-free asset with the market proxy alone. In other words, the Gibbons, Ross and Shanken statistic tests whether the market proxy is the tangency portfolio in the set of portfolios that can be constructed by combining the market portfolio with the specific assets used as dependent variables in the time-series regressions.

Enlightened by this insight of Gibbons, Ross and Shanken (1989), one can see a similar interpretation of the cross-section regression test of whether market betas suffice to explain expected returns. In this case, the test is whether the additional explanatory variables in a cross-section regression identify patterns in the returns on the left-hand-side assets that are not explained by the assets' market betas. This amounts to testing whether the market proxy is on the minimum variance frontier that can be constructed using the market proxy and the left-hand-side assets included in the tests.

An important lesson from this discussion is that time-series and cross-section regressions do not, strictly speaking, test the CAPM. What is literally tested is whether a specific proxy for the market portfolio (typically a portfolio of U.S. common stocks) is efficient in the set of portfolios that can be constructed from it and the left-hand-side assets used in the test. One might conclude from this that the CAPM has never been tested, and prospects for testing it are not good because 1) the set of left-hand-side assets does not include all marketable assets, and 2) data for the true market portfolio of all assets are likely beyond reach (Roll, 1977; more on this later). But this criticism can be leveled at tests of any economic model when the tests are less than exhaustive or when they use proxies for the variables called for by the model.

The bottom line from the early cross-section regression tests of the CAPM, such as Fama and MacBeth (1973), and the early time-series regression tests, like Gibbons (1982) and Stambaugh (1982), is that standard market proxies seem to be on the minimum variance frontier. That is, the central predictions of the Black version of the CAPM, that market betas suffice to explain expected returns and that the risk premium for beta is positive, seem to hold. But the more specific prediction of the Sharpe-Lintner CAPM that the premium per unit of beta is the expected market return minus the risk-free interest rate is consistently rejected.

The success of the Black version of the CAPM in early tests produced a consensus that the model is a good description of expected returns. These early results, coupled with the model's simplicity and intuitive appeal, pushed the CAPM to the forefront of finance.

Recent Tests

Starting in the late 1970s, empirical work appears that challenges even the Black version of the CAPM. Specifically, evidence mounts that much of the variation in expected return is unrelated to market beta.

The first blow is Basu's (1977) evidence that when common stocks are sorted on earnings-price ratios, future returns on high E/P stocks are higher than predicted by the CAPM. Banz (1981) documents a size effect: when stocks are sorted on market capitalization (price times shares outstanding), average returns on small stocks are higher than predicted by the CAPM. Bhandari (1988) finds that high debt-equity ratios (book value of debt over the market value of equity, a measure of leverage) are associated with returns that are too high relative to their market betas.

Finally, Statman (1980) and Rosenberg, Reid and Lanstein (1985) document that stocks with high book-to-market equity ratios (B/M, the ratio of the book value of a common stock to its market value) have high average returns that are not captured by their betas.

There is a theme in the contradictions of the CAPM summarized above. Ratios involving stock prices have information about expected returns missed by market betas. On reflection, this is not surprising. A stock's price depends not only on the expected cash flows it will provide, but also on the expected returns that discount expected cash flows back to the present. Thus, in principle, the cross-section of prices has information about the cross-section of expected returns. (A high expected return implies a high discount rate and a low price.) The cross-section of stock prices is, however, arbitrarily affected by differences in scale (or units). But with a judicious choice of scaling variable X, the ratio X/P can reveal differences in the cross-section of expected stock returns. Such ratios are thus prime candidates to expose shortcomings of asset pricing models—in the case of the CAPM, shortcomings of the prediction that market betas suffice to explain expected returns (Ball, 1978). The contradictions of the CAPM summarized above suggest that earnings-price, debt-equity and book-to-market ratios indeed play this role.

Fama and French (1992) update and synthesize the evidence on the empirical failures of the CAPM. Using the cross-section regression approach, they confirm that size, earnings-price, debt-equity and book-to-market ratios add to the explanation of expected stock returns provided by market beta. Fama and French (1996) reach the same conclusion using the time-series regression approach applied to portfolios of stocks sorted on price ratios. They also find that different price ratios have much the same information about expected returns. This is not surprising given that price is the common driving force in the price ratios, and the numerators are just scaling variables used to extract the information in price about expected returns.

Fama and French (1992) also confirm the evidence (Reinganum, 1981; Stambaugh, 1982; Lakonishok and Shapiro, 1986) that the relation between average return and beta for common stocks is even flatter after the sample periods used in the early empirical work on the CAPM. The estimate of the beta premium is, however, clouded by statistical uncertainty (a large standard error). Kothari, Shanken and Sloan (1995) try to resuscitate the Sharpe-Lintner CAPM by arguing that the weak relation between average return and beta is just a chance result. But the strong evidence that other variables capture variation in expected return missed by beta makes this argument irrelevant. If betas do not suffice to explain expected returns, the market portfolio is not efficient, and the CAPM is dead in its tracks. Evidence on the size of the market premium can neither save the model nor further doom it.

The synthesis of the evidence on the empirical problems of the CAPM provided by Fama and French (1992) serves as a catalyst, marking the point when it is generally acknowledged that the CAPM has potentially fatal problems. Research then turns to explanations.

One possibility is that the CAPM's problems are spurious, the result of data dredging—publication-hungry researchers scouring the data and unearthing contradictions that occur in specific samples as a result of chance. A standard response to this concern is to test for similar findings in other samples. Chan, Hamao and Lakonishok (1991) find a strong relation between book-to-market equity (B/M) and average return for Japanese stocks. Capaul, Rowley and Sharpe (1993) observe a similar B/M effect in four European stock markets and in Japan. Fama and French (1998) find that the price ratios that produce problems for the CAPM in U.S. data show up in the same way in the stock returns of twelve non-U.S. major markets, and they are present in emerging market returns. This evidence suggests that the contradictions of the CAPM associated with price ratios are not sample specific.

Explanations: Irrational Pricing or Risk

Among those who conclude that the empirical failures of the CAPM are fatal, two stories emerge. On one side are the behavioralists. Their view is based on evidence that stocks with high ratios of book value to market price are typically firms that have fallen on bad times, while low B/M is associated with growth firms (Lakonishok, Shleifer and Vishny, 1994; Fama and French, 1995). The behavioralists argue that sorting firms on book-to-market ratios exposes investor overreaction to good and bad times. Investors overextrapolate past performance, resulting in stock prices that are too high for growth (low B/M) firms and too low for distressed (high B/M, so-called value) firms. When the overreaction is eventually corrected, the result is high returns for value stocks and low returns for growth stocks. Proponents of this view include DeBondt and Thaler (1987), Lakonishok, Shleifer and Vishny (1994) and Haugen (1995).

The second story for explaining the empirical contradictions of the CAPM is that they point to the need for a more complicated asset pricing model. The CAPM is based on many unrealistic assumptions. For example, the assumption that investors care only about the mean and variance of one-period portfolio returns is extreme. It is reasonable that investors also care about how their portfolio return covaries with labor income and future investment opportunities, so a portfolio's return variance misses important dimensions of risk. If so, market beta is not a complete description of an asset's risk, and we should not be surprised to find that differences in expected return are not completely explained by differences in beta. In this view, the search should turn to asset pricing models that do a better job explaining average returns.

Merton's (1973) intertemporal capital asset pricing model (ICAPM) is a natural extension of the CAPM. The ICAPM begins with a different assumption about investor objectives. In the CAPM, investors care only about the wealth their portfolio produces at the end of the current period. In the ICAPM, investors are concerned not only with their end-of-period payoff, but also with the opportunities they will have to consume or invest the payoff. Thus, when choosing a portfolio at time $t-1$, ICAPM investors consider how their wealth at t might vary with future *state variables,* including labor income, the prices of consumption goods and the nature of portfolio opportunities at t, and expectations about the labor income, consumption and investment opportunities to be available after t.

Like CAPM investors, ICAPM investors prefer high expected return and low return variance. But ICAPM investors are also concerned with the covariances of portfolio returns with state variables. As a result, optimal portfolios are "multifactor efficient," which means they have the largest possible expected returns, given their return variances and the covariances of their returns with the relevant state variables.

Fama (1996) shows that the ICAPM generalizes the logic of the CAPM. That is, if there is risk-free borrowing and lending or if short sales of risky assets are allowed, market clearing prices imply that the market portfolio is multifactor efficient. Moreover, multifactor efficiency implies a relation between expected return and beta risks, but it requires additional betas, along with a market beta, to explain expected returns.

An ideal implementation of the ICAPM would specify the state variables that affect expected returns. Fama and French (1993) take a more indirect approach, perhaps more in the spirit of Ross's (1976) arbitrage pricing theory. They argue that though size and book-to-market equity are not themselves state variables, the higher average returns on small stocks and high book-to-market stocks reflect unidentified state variables that produce undiversifiable risks (covariances) in returns that are not captured by the market return and are priced separately from market betas. In support of this claim, they show that the returns on the stocks of small firms covary more with one another than with returns on the stocks of large firms, and returns on high book-to-market (value) stocks covary more with one another than with returns on low book-to-market (growth) stocks. Fama and French (1995) show that there are similar size and book-to-market patterns in the covariation of fundamentals like earnings and sales.

Based on this evidence, Fama and French (1993, 1996) propose a three-factor model for expected returns,

$$\text{(Three-Factor Model)} \quad E(R_{it}) - R_{ft} = \beta_{iM}[E(R_{Mt}) - R_{ft}] + \beta_{is}E(SMB_t) + \beta_{ih}E(HML_t).$$

In this equation, SMB_t (small minus big) is the difference between the returns on diversified portfolios of small and big stocks, HML_t (high minus low) is the difference between the returns on diversified portfolios of high and low B/M stocks, and the betas are slopes in the multiple regression of $R_{it} - R_{ft}$ on $R_{Mt} - R_{ft}$, SMB_t and HML_t.

For perspective, the average value of the market premium $R_{Mt} - R_{ft}$ for 1927–2003 is 8.3 percent per year, which is 3.5 standard errors from zero. The average values of SMB_t, and HML_t are 3.6 percent and 5.0 percent per year, and they are 2.1 and 3.1 standard errors from zero. All three premiums are volatile, with annual standard deviations of 21.0 percent ($R_{Mt} - R_{ft}$), 14.6 percent (SMB_t) and 14.2 percent (HML_t) per year. Although the average values of the premiums are large, high volatility implies substantial uncertainty about the true expected premiums.

One implication of the expected return equation of the three-factor model is that the intercept α_i in the time-series regression,

$$R_{it} - R_{ft} = \alpha_i + \beta_{iM}(R_{Mt} - R_{ft}) + \beta_{is}SMB_t + \beta_{ih}HML_t + \varepsilon_{it},$$

is zero for all assets i. Using this criterion, Fama and French (1993, 1996) find that the model captures much of the variation in average return for portfolios formed on size, book-to-market equity and other price ratios that cause problems for the CAPM. Fama and French (1998) show that an international version of the model performs better than an international CAPM in describing average returns on portfolios formed on scaled price variables for stocks in 13 major markets.

The three-factor model is now widely used in empirical research that requires a model of expected returns. Estimates of α_i from the time-series regression above are used to calibrate how rapidly stock prices respond to new information (for example, Loughran and Ritter, 1995; Mitchell and Stafford, 2000). They are also used to measure the special information of portfolio managers, for example, in Carhart's (1997) study of mutual fund performance. Among practitioners like Ibbotson Associates, the model is offered as an alternative to the CAPM for estimating the cost of equity capital.

From a theoretical perspective, the main shortcoming of the three-factor model is its empirical motivation. The small-minus-big (SMB) and high-minus-low (HML) explanatory returns are not motivated by predictions about state variables of concern to investors. Instead they are brute force constructs meant to capture the patterns uncovered by previous work on how average stock returns vary with size and the book-to-market equity ratio.

But this concern is not fatal. The ICAPM does not require that the additional portfolios used along with the market portfolio to explain expected returns "mimic" the relevant state variables. In both the ICAPM and the arbitrage pricing theory, it suffices that the additional portfolios are well diversified (in the terminology of Fama, 1996, they are multifactor minimum variance) and that they are sufficiently different from the market portfolio to capture covariation in returns and variation in expected returns missed by the market portfolio. Thus, adding diversified portfolios that capture covariation in returns and variation in average returns left unexplained by the market is in the spirit of both the ICAPM and the Ross's arbitrage pricing theory.

The behavioralists are not impressed by the evidence for a risk-based explanation of the failures of the CAPM. They typically concede that the three-factor model captures covariation in returns missed by the market return and that it picks up much of the size and value effects in average returns left unexplained by the CAPM. But their view is that the average return premium associated with the model's book-to-market factor—which does the heavy lifting in the improvements to the CAPM—is itself the result of investor overreaction that happens to be correlated across firms in a way that just looks like a risk story. In short, in the behavioral view, the market tries to set CAPM prices, and violations of the CAPM are due to mispricing.

The conflict between the behavioral irrational pricing story and the rational risk story for the empirical failures of the CAPM leaves us at a timeworn impasse. Fama (1970) emphasizes that the hypothesis that prices properly reflect available information must be tested in the context of a model of expected returns, like the CAPM. Intuitively, to test whether prices are rational, one must take a stand on what the market is trying to do in setting prices—that is, what is risk and what is the relation between expected return and risk? When tests reject the CAPM, one cannot say whether the problem is its assumption that prices are rational (the behavioral view) or violations of other assumptions that are also necessary to produce the CAPM (our position).

Fortunately, for some applications, the way one uses the three-factor model does not depend on one's view about whether its average return premiums are the rational result of underlying state variable risks, the result of irrational investor behavior or sample specific results of chance. For example, when measuring the response of stock prices to new information or when evaluating the performance of managed portfolios, one wants to account for known patterns in returns and average returns for the period examined, whatever their source. Similarly, when estimating the cost of equity capital, one might be unconcerned with whether expected return premiums are rational or irrational since they are in either case part of the opportunity cost of equity capital (Stein, 1996). But the cost of capital is forward looking, so if the premiums are sample specific they are irrelevant.

The three-factor model is hardly a panacea. Its most serious problem is the momentum effect of Jegadeesh and Titman (1993). Stocks that do well relative to the market over the last three to twelve months tend to continue to do well for the next few months, and stocks that do poorly continue to do poorly. This momentum effect is distinct from the value effect captured by book-to-market equity and other price ratios. Moreover, the momentum effect is left unexplained by the three-factor model, as well as by the CAPM. Following Carhart (1997), one response is to add a momentum factor (the difference between the returns on diversified portfolios of short-term winners and losers) to the three-factor model. This step is again legitimate in applications where the goal is to abstract from known patterns in average returns to uncover information-specific or manager-specific effects. But since the momentum effect is short-lived, it is largely irrelevant for estimates of the cost of equity capital.

Another strand of research points to problems in both the three-factor model and the CAPM. Frankel and Lee (1998), Dechow, Hutton and Sloan (1999), Piotroski (2000) and others show that in portfolios formed on price ratios like book-to-market equity, stocks with higher expected cash flows have higher average returns that are not captured by the three-factor model or the CAPM. The authors interpret their results as evidence that stock prices are irrational, in the sense that they do not reflect available information about expected profitability.

In truth, however, one can't tell whether the problem is bad pricing or a bad asset pricing model. A stock's price can always be expressed as the present value of expected future cash flows discounted at the expected return on the stock (Campbell and Shiller, 1989; Vuolteenaho, 2002). It follows that if two stocks have the same price, the one with higher expected cash flows must have a higher expected return. This holds true whether pricing is rational or irrational. Thus, when one observes a positive relation between expected cash flows and expected returns that is left unexplained by the CAPM or the three-factor model, one can't tell whether it is the result of irrational pricing or a misspecified asset pricing model.

The Market Proxy Problem

Roll (1977) argues that the CAPM has never been tested and probably never will be. The problem is that the market portfolio at the heart of the model is theoretically and empirically elusive. It is not theoretically clear which assets (for example, human capital) can legitimately be excluded from the market portfolio, and data availability substantially limits the assets that are included. As a result, tests of the CAPM are forced to use proxies for the market portfolio, in effect testing whether the proxies are on the minimum variance frontier. Roll argues that because the tests use proxies, not the true market portfolio, we learn nothing about the CAPM.

We are more pragmatic. The relation between expected return and market beta of the CAPM is just the minimum variance condition that holds in any efficient portfolio, applied to the market portfolio. Thus, if we can find a market proxy that is on the minimum variance frontier, it can be used to describe differences in expected returns, and we would be happy to use it for this purpose. The strong rejections of the CAPM described above, however, say that researchers have not uncovered a reasonable market proxy that is close to the minimum variance frontier. If researchers are constrained to reasonable proxies, we doubt they ever will.

Our pessimism is fueled by several empirical results. Stambaugh (1982) tests the CAPM using a range of market portfolios that include, in addition to U.S. common stocks, corporate and government bonds, preferred stocks, real estate and other consumer durables. He finds that tests of the CAPM are not sensitive to expanding the market proxy beyond common stocks, basically because the volatility of expanded market returns is dominated by the volatility of stock returns.

One need not be convinced by Stambaugh's (1982) results since his market proxies are limited to U.S. assets. If international capital markets are open and asset prices conform to an international version of the CAPM, the market portfolio should include international assets. Fama and French (1998) find, however, that betas for a global stock market portfolio cannot explain the high average returns observed around the world on stocks with high book-to-market or high earnings-price ratios.

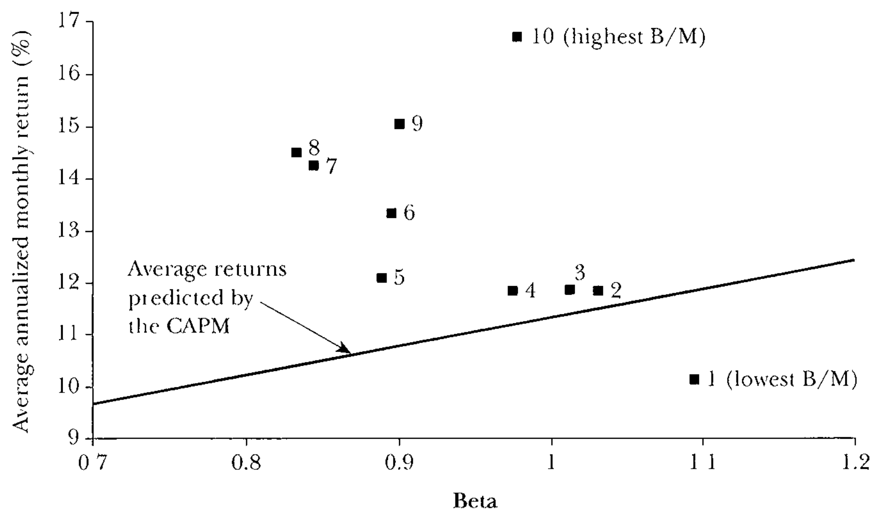
A major problem for the CAPM is that portfolios formed by sorting stocks on price ratios produce a wide range of average returns, but the average returns are not positively related to market betas (Lakonishok, Shleifer and Vishny, 1994; Fama and French, 1996, 1998). The problem is illustrated in Figure 3, which shows average returns and betas (calculated with respect to the CRSP value-weight portfolio of NYSE, AMEX and NASDAQ stocks) for July 1963 to December 2003 for ten portfolios of U.S. stocks formed annually on sorted values of the book-to-market equity ratio (B/M).[6]

Average returns on the B/M portfolios increase almost monotonically, from 10.1 percent per year for the lowest B/M group (portfolio 1) to an impressive 16.7 percent for the highest (portfolio 10). But the positive relation between beta and average return predicted by the CAPM is notably absent. For example, the portfolio with the lowest book-to-market ratio has the highest beta but the lowest average return. The estimated beta for the portfolio with the highest book-to-market ratio and the highest average return is only 0.98. With an average annualized value of the riskfree interest rate, R_f, of 5.8 percent and an average annualized market premium, $R_M - R_f$, of 11.3 percent, the Sharpe-Lintner CAPM predicts an average return of 11.8 percent for the lowest B/M portfolio and 11.2 percent for the highest, far from the observed values, 10.1 and 16.7 percent. For the Sharpe-Lintner model to "work" on these portfolios, their market betas must change dramatically, from 1.09 to 0.78 for the lowest B/M portfolio and from 0.98 to 1.98 for the highest. We judge it unlikely that alternative proxies for the market portfolio will produce betas and a market premium that can explain the average returns on these portfolios.

It is always possible that researchers will redeem the CAPM by finding a reasonable proxy for the market portfolio that is on the minimum variance frontier. We emphasize, however, that this possibility cannot be used to justify the way the CAPM is currently applied. The problem is that applications typically use the same

[6] Stock return data are from CRSP, and book equity data are from Compustat and the Moody's Industrials, Transportation, Utilities and Financials manuals. Stocks are allocated to ten portfolios at the end of June of each year t (1963 to 2003) using the ratio of book equity for the fiscal year ending in calendar year $t - 1$, divided by market equity at the end of December of $t - 1$. Book equity is the book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, we use the redemption, liquidation or par value (in that order) to estimate the book value of preferred stock. Stockholders' equity is the value reported by Moody's or Compustat, if it is available. If not, we measure stockholders' equity as the book value of common equity plus the par value of preferred stock or the book value of assets minus total liabilities (in that order). The portfolios for year t include NYSE (1963–2003), AMEX (1963–2003) and NASDAQ (1972–2003) stocks with positive book equity in $t - 1$ and market equity (from CRSP) for December of $t - 1$ and June of t. The portfolios exclude securities CRSP does not classify as ordinary common equity. The breakpoints for year t use only securities that are on the NYSE in June of year t.

Figure 3
Average Annualized Monthly Return versus Beta for Value Weight Portfolios Formed on B/M, 1963–2003



market proxies, like the value-weight portfolio of U.S. stocks, that lead to rejections of the model in empirical tests. The contradictions of the CAPM observed when such proxies are used in tests of the model show up as bad estimates of expected returns in applications; for example, estimates of the cost of equity capital that are too low (relative to historical average returns) for small stocks and for stocks with high book-to-market equity ratios. In short, if a market proxy does not work in tests of the CAPM, it does not work in applications.

Conclusions

The version of the CAPM developed by Sharpe (1964) and Lintner (1965) has never been an empirical success. In the early empirical work, the Black (1972) version of the model, which can accommodate a flatter tradeoff of average return for market beta, has some success. But in the late 1970s, research begins to uncover variables like size, various price ratios and momentum that add to the explanation of average returns provided by beta. The problems are serious enough to invalidate most applications of the CAPM.

For example, finance textbooks often recommend using the Sharpe-Lintner CAPM risk-return relation to estimate the cost of equity capital. The prescription is to estimate a stock's market beta and combine it with the risk-free interest rate and the average market risk premium to produce an estimate of the cost of equity. The typical market portfolio in these exercises includes just U.S. common stocks. But empirical work, old and new, tells us that the relation between beta and average return is flatter than predicted by the Sharpe-Lintner version of the CAPM. As a result, CAPM estimates of the cost of equity for high beta stocks are too high (relative to historical average returns) and estimates for low beta stocks are too low (Friend and Blume, 1970). Similarly, if the high average returns on value stocks (with high book-to-market ratios) imply high expected returns, CAPM cost of equity estimates for such stocks are too low.[7]

The CAPM is also often used to measure the performance of mutual funds and other managed portfolios. The approach, dating to Jensen (1968), is to estimate the CAPM time-series regression for a portfolio and use the intercept (Jensen's alpha) to measure abnormal performance. The problem is that, because of the empirical failings of the CAPM, even passively managed stock portfolios produce abnormal returns if their investment strategies involve tilts toward CAPM problems (Elton, Gruber, Das and Hlavka, 1993). For example, funds that concentrate on low beta stocks, small stocks or value stocks will tend to produce positive abnormal returns relative to the predictions of the Sharpe-Lintner CAPM, even when the fund managers have no special talent for picking winners.

The CAPM, like Markowitz's (1952, 1959) portfolio model on which it is built, is nevertheless a theoretical tour de force. We continue to teach the CAPM as an introduction to the fundamental concepts of portfolio theory and asset pricing, to be built on by more complicated models like Merton's (1973) ICAPM. But we also warn students that despite its seductive simplicity, the CAPM's empirical problems probably invalidate its use in applications.

■ *We gratefully acknowledge the comments of John Cochrane, George Constantinides, Richard Leftwich, Andrei Shleifer, René Stulz and Timothy Taylor.*

[7] The problems are compounded by the large standard errors of estimates of the market premium and of betas for individual stocks, which probably suffice to make CAPM estimates of the cost of equity rather meaningless, even if the CAPM holds (Fama and French, 1997; Pastor and Stambaugh, 1999). For example, using the U.S. Treasury bill rate as the risk-free interest rate and the CRSP value-weight portfolio of publicly traded U.S. common stocks, the average value of the equity premium $R_{Mt} - R_{ft}$ for 1927–2003 is 8.3 percent per year, with a standard error of 2.4 percent. The two standard error range thus runs from 3.5 percent to 13.1 percent, which is sufficient to make most projects appear either profitable or unprofitable. This problem is, however, hardly special to the CAPM. For example, expected returns in all versions of Merton's (1973) ICAPM include a market beta and the expected market premium. Also, as noted earlier the expected values of the size and book-to-market premiums in the Fama-French three-factor model are also estimated with substantial error.

References

Ball, Ray. 1978. "Anomalies in Relationships Between Securities' Yields and Yield-Surrogates." *Journal of Financial Economics.* 6:2, pp. 103–26.

Banz, Rolf W. 1981 "The Relationship Between Return and Market Value of Common Stocks." *Journal of Financial Economics.* 9:1, pp. 3–18

Basu, Sanjay. 1977. "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis." *Journal of Finance.* 12:3, pp. 129–56

Bhandari, Laxmi Chand. 1988. "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence." *Journal of Finance.* 43.2, pp. 507–28.

Black, Fischer. 1972. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business.* 45:3, pp. 444–54.

Black, Fischer, Michael C. Jensen and Myron Scholes. 1972. "The Capital Asset Pricing Model: Some Empirical Tests," in *Studies in the Theory of Capital Markets.* Michael C. Jensen, ed. New York: Praeger, pp. 79–121.

Blume, Marshall. 1970. "Portfolio Theory: A Step Towards its Practical Application." *Journal of Business.* 43:2, pp. 152–74.

Blume, Marshall and Irwin Friend. 1973. "A New Look at the Capital Asset Pricing Model." *Journal of Finance.* 28:1, pp 19–33.

Campbell, John Y. and Robert J. Shiller. 1989. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies* 1:3, pp 195–228.

Capaul, Carlo, Ian Rowley and William F. Sharpe. 1993. "International Value and Growth Stock Returns." *Financial Analysts Journal.* January/February, 49, pp 27–36

Carhart, Mark M. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance.* 52:1, pp. 57–82.

Chan, Louis K.C., Yasushi Hamao and Josef Lakonishok. 1991. "Fundamentals and Stock Returns in Japan." *Journal of Finance.* 46:5, pp. 1739–789.

DeBondt, Werner F. M. and Richard H. Thaler. 1987. "Further Evidence on Investor Overreaction and Stock Market Seasonality." *Journal of Finance.* 42:3, pp. 557–81.

Dechow, Patricia M., Amy P. Hutton and Richard G. Sloan. 1999. "An Empirical Assessment of the Residual Income Valuation Model." *Journal of Accounting and Economics.* 26:1, pp. 1–34.

Douglas, George W. 1968. *Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency* Ann Arbor, Michigan: University Microfilms, Inc.

Elton, Edwin J., Martin J. Gruber, Sanjiv Das and Matt Hlavka. 1993. "Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios." *Review of Financial Studies.* 6:1, pp. 1–22.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25:2, pp. 383–417.

Fama, Eugene F. 1996. "Multifactor Portfolio Efficiency and Multifactor Asset Pricing." *Journal of Financial and Quantitative Analysis.* 31:4, pp. 441–65.

Fama, Eugene F. and Kenneth R. French. 1992 "The Cross-Section of Expected Stock Returns." *Journal of Finance.* 47:2, pp. 427–65.

Fama, Eugene F. and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds" *Journal of Financial Economics.* 33.1, pp. 3–56.

Fama, Eugene F. and Kenneth R. French. 1995. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Finance.* 50:1, pp. 131–55.

Fama, Eugene F. and Kenneth R. French. 1996. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance.* 51.1, pp. 55–84.

Fama, Eugene F. and Kenneth R. French. 1997 "Industry Costs of Equity." *Journal of Financial Economics.* 43:2 pp. 153–93.

Fama, Eugene F. and Kenneth R. French. 1998. "Value Versus Growth: The International Evidence." *Journal of Finance.* 53.6, pp. 1975–999.

Fama, Eugene F. and James D. MacBeth. 1973. "Risk, Return, and Equilibrium Empirical Tests." *Journal of Political Economy.* 81:3, pp. 607–36.

Frankel, Richard and Charles M.C. Lee. 1998. "Accounting Valuation, Market Expectation, and Cross-Sectional Stock Returns." *Journal of Accounting and Economics.* 25.3 pp. 283–319.

Friend, Irwin and Marshall Blume. 1970. "Measurement of Portfolio Performance under Uncertainty" *American Economic Review.* 60:4, pp. 607–36.

Gibbons, Michael R. 1982. "Multivariate Tests of Financial Models: A New Approach." *Journal of Financial Economics.* 10 1, pp 3–27.

Gibbons, Michael R., Stephen A. Ross and Jay Shanken. 1989. "A Test of the Efficiency of a Given Portfolio." *Econometrica.* 57.5, pp 1121–152.

Haugen, Robert. 1995. *The New Finance: The*